



Regionalización de Curvas de Duración de Caudales en Colombia

Carlos José Gaviria Arbeláez

Universidad Nacional de Colombia
Facultad de Minas, Departamento de Geociencias y Medio Ambiente
Medellín, Colombia
2018

Regionalización de Curvas de Duración de Caudales en Colombia

Carlos José Gaviria Arbeláez

Tesis presentada como requisito para optar al título de:

Magister en Ingeniería de los Recursos Hidráulicos

Director:

Luis Fernando Carvajal Serna, IC MSc.

Línea de Investigación:

Sistemas Hidrológicos

Grupo de Investigación:

Posgrado en Aprovechamiento de Recursos Hidráulicos

Universidad Nacional de Colombia

Facultad de Minas, Departamento de Geociencias y Medio Ambiente

Medellín, Colombia

2018

Puede que lo que hacemos no traiga siempre
felicidad, pero si no hacemos nada, no habrá
felicidad.

Albert Camus

Agradecimientos

A mi familia, en especial a mis padres que con paciencia y ternura me enseñaron el valor de aprender, a mis hermanos por su grata compañía y a mi pareja, Angélica Salgado, por todo el amor, el cariño sincero y el apoyo incondicional.

A mi director Luis Fernando Carvajal Serna, por el tiempo dedicado, la asesoría brindada y por nunca negarse a realizar aportes desde su experiencia y conocimiento del tema. Al Área Curricular en Medio Ambiente, al Departamento de Geociencias y Medio Ambiente y a la Facultad de Minas por la beca otorgada, así como a los profesores Jaime Ignacio Vélez Upegui, Mauricio Toro Botero, Jorge Julián Vélez Upegui, Lilian Posada García y Andrés Osorio Arias, por haberme brindado la oportunidad de trabajar con cada uno de ellos.

A los institutos IDEAM, IGAC y SGC por dejar disponible la información necesaria para el desarrollo de este trabajo, así como a los profesores Jaime Ignacio Vélez y María Victoria Vélez por aceptar ser los jurados, dedicar el tiempo a la revisión y realizar comentarios constructivos del trabajo. A los profesores Alejandro Jaramillo Moreno, de Mecánica de Fluidos, y Julián David Rojo Hernández, de Hidrología, por despertar, desde mi pregrado, el interés que derivó en la inclinación hacia esta rama de la ingeniería. Agradezco también al Colegio Calasanz y a la Universidad Nacional por la excelente formación académica recibida.

A Alejandro García Ramírez, Sebastián Gómez Ríos, Juan David Parra Jiménez, Santiago Cataño Álvarez, Mateo Gutiérrez Uribe y Manuel Cárdenas Quintero por el interés y porque nunca dudaron en brindarme una mano y consejos cuando los necesité. A mis amigos y compañeros de posgrado; Edward Julián Sánchez Trujillo, Stephany Córdoba Yepes, Claudia Bedoya Cardona, Manuel Coy Pertuz, Felipe Arbeláez Salazar, Adrián Perpiñán Guerra, Beatriz Martínez Contreras, Mónica Bonilla Rodríguez, Laura Rodríguez Martínez, Sandra Salamanca Jiménez, Mateo Bonilla Ramírez, Alejandro Cáceres Euse, Eliana Bohórquez Bedoya, Diana Arroyave, Adriana Puello, Enovaldo Herrera, Maximiliano Arredondo, Gabriel Rosado, y a muchos otros, porque contribuyeron a que esta experiencia fuese algo verdaderamente ameno y agradable.

Resumen

Es de gran importancia conocer el comportamiento de los caudales de una cuenca hidrológica para la adecuada planeación y manejo del recurso hídrico, sin embargo, en Colombia existen muchas corrientes sin ningún tipo de registro o con información escasa, lo que genera la necesidad de implementar modelos hidrológicos que permitan llevar a cabo la estimación de caudales. El régimen hidrológico puede resumirse gráficamente con la curva de duración de caudales CDC (magnitud vs frecuencia o porcentaje de excedencia). Se utilizan series diarias de caudales provenientes de 655 estaciones, además de información geológica, topográfica y climática disponible para todo el país, para definir regiones hidrológicamente homogéneas. Mediante un análisis de clúster utilizando el algoritmo *K-means*, se delimitaron 15 regiones hidrológicas en el país, las cuales presentaron una alta heterogeneidad espacial. Se utilizaron regresiones lineales múltiples para estimar caudales característicos adimensionales (Q_p/\bar{Q}) en función de los atributos extraídos de cada cuenca. Se obtiene una serie de ecuaciones que permiten reconstruir curvas de duración de caudales adimensionales sintéticas para cada región hidrológica, presentándose valores de R^2 promedio entre 0.5 y 0.9, las comparaciones con las CDC observadas arrojaron errores porcentuales promedio de aproximadamente 9% en descargas medias, 30% en máximas y 50% en mínimas. Se analizaron cambios temporales en algunas CDC mediante la aplicación del test no paramétrico de *Mann-Kendall*, la pendiente de Sen y la pendiente de la regresión lineal de la serie, se observan tendencias crecientes significativas en las regiones del Chocó y Bajo cauca, mientras que para tendencias decrecientes no se observa ningún patrón.

Palabras clave: Regionalización, Curva de duración de caudales, Regiones hidrológicas, K-means clustering, Regresión lineal múltiple, Análisis no estacionario.

Abstract

For an appropriate water resource planning and management, it is important to know the behavior of the streamflow in hydrological basins, however, in Colombia there are many ungauged catchments, which generates the need of running hydrological models. The flow regime can be graphically summarized by the flow duration curve FDC (magnitude vs. frequency). Daily flow time series from 655 gauging stations are used to define homogenous hydrological regions taking account geological, topographic and climatic information available for the whole country. 15 hydrological regions were delimited by cluster analysis using *K-means* algorithm, these regions exhibit a high spatial heterogeneity. Multiple linear regressions were used in order to estimate different characteristic dimensionless flows (Q_p/\bar{Q}) as a function of each basin's attributes. As a result, we obtained a set of equations which allow the reconstruction of simulated dimensionless duration curves for each region, regression R^2 values between 0.5 and 0.9 were obtained. Simulated FDC were compared with observed values getting average percentage errors of approximately 9% for mean flows, 30% for maximum flow rates and 50% in minimums. Long term temporal changes are analyzed in several FDC by the mean of the application of non-parametric *Mann-Kendall* test, the Sen Slope and the slope of the linear regression of the time series. Significant increasing trends are observed in the regions of Chocó and Bajo Cauca, while decreasing trends did not present any particular pattern.

Keywords: Regionalization, Flow duration curves, Hydrological regions, K-means clustering, Lineal multiple regression, Non stationary analysis.

Contenido

	Pág.
Resumen.....	IX
Lista de figuras	XIII
Lista de tablas	XVII
1 Introducción	1
1.1 Generalidades	1
1.2 Justificación.....	5
1.3 Antecedentes	7
1.3.1 Definición de regiones hidrológicas	7
1.3.2 Estimación de curvas de duración regionales.....	8
1.3.3 Análisis no estacionario	10
1.4 Objetivos	13
1.4.1 Objetivo general	13
1.4.2 Objetivos específicos	13
1.5 Metodología.....	13
2 Datos.....	15
2.1 Información	15
2.1.1 Caudales	15
2.1.2 Modelo digital de elevación	16
2.1.3 Mapa de coberturas	17
2.1.4 Mapa geológico de Colombia	18
2.1.5 Campos de precipitación mensual de Colombia.....	20
2.1.6 Evapotranspiración potencial	22
2.1.7 Temperatura media en superficie.....	23
2.2 Análisis y procesamiento de la información.....	24
2.2.1 Tratamiento de las series de caudales.....	24
2.2.2 Procesamiento y manejo del modelo digital de elevación	26
2.2.3 Cuencas calibradas y cuencas sin calibrar	28
2.3 Análisis de calidad de la información	31

3	Definición de Regiones Hidrológicas de Colombia	33
3.1	Introducción	33
3.2	Metodología.....	34
3.3	Modelo de agrupación de cuencas	37
3.3.1	Selección de número de grupos.....	40
3.4	Aplicación de clúster y resultados	43
3.4.1	Agrupación de cuencas sin calibrar	45
3.5	Análisis de resultados.....	46
4	Regionalización de Curvas de Duración de Caudales	53
4.1	Introducción	53
4.2	Modelo de estimación de caudales.....	54
4.2.1	Selección de percentiles característicos	54
4.2.2	Selección de estaciones para validación del modelo	54
4.2.3	Insumos para la calibración del modelo.....	55
4.2.4	Regresión lineal múltiple	56
4.2.5	Resultados	57
4.3	Análisis de resultados.....	69
5	Análisis No Estacionario de la Curva de Duración de Caudales en Colombia	73
5.1	Introducción y metodología.....	73
5.2	Selección de series temporales	75
5.3	Resultados.....	79
5.4	Análisis de resultados.....	92
6	Conclusiones y Recomendaciones	95
6.1	Conclusiones.....	95
6.2	Recomendaciones y trabajo futuro	97
A.	Resumen de Resultados Matriz de Correlaciones	101
B.	Centroides geométricos de cada clúster.....	107
C.	Resultados de Regresiones Lineales Múltiples	109
D.	Dispersiones Atributos vs Adre	127
	Bibliografía	129

Lista de figuras

	Pág.
Figura 2-1. Estaciones de caudales diarios en Colombia, red nacional de mediciones del IDEAM.....	16
Figura 2-2. Modelo de Elevación Digital de Colombia - resolución 90 m, fuente: <i>SRTM</i> . 17	17
Figura 2-3. Mapa de coberturas de Colombia, escala: 1:100.000, fuente: <i>IGAC</i>	18
Figura 2-4. Mapa geológico de Colombia, escala 1:100.000, fuente: <i>IGAC</i> - <i>SGC</i>	19
Figura 2-5. Mapa de fallas geológicas de Colombia. Escala: 1:100.000, fuente: <i>IGAC</i> - <i>SGC</i>	20
Figura 2-6. Precipitación media mensual de Colombia 1975 – 2006, fuente: Hurtado y Mesa (2014).	21
Figura 2-7. Precipitación máxima mensual de Colombia 1975 – 2006, fuente: Hurtado y Mesa (2014).....	22
Figura 2-8. Evapotranspiración potencial anual de Colombia.	23
Figura 2-9. Temperatura media del aire en superficie en °C para Colombia.....	24
Figura 2-10. Estación ubicada en brazo.	25
Figura 2-11. Estación ubicada en canal artificial.....	25
Figura 2-12. Estaciones de medición de caudal definitivas a utilizar, 655 en total.	26
Figura 2-13. Ejemplo de contraste entre redes en zona montañosa.....	27
Figura 2-14. Ejemplo de contraste entre redes en zona plana.	27
Figura 2-15. Mapa de cuencas calibradas trazadas.....	29
Figura 2-16. Mapa de cuencas sin calibrar en el territorio colombiano. Fuente: (Salazar-Holguín, 2013).	30
Figura 2-17. Caudal medio a largo plazo vs área de drenaje en las 655 series de caudales.	32
Figura 3-1. CDC real estación 11037010.	35
Figura 3-2. CDC real estación 21047010.	35
Figura 3-3. CDC adim. Estación 11037010.	36
Figura 3-4. CDC adim. Estación 21047010.	36
Figura 3-5. CDC Normalizada estación 11037010.	36

Figura 3-6. CDC Normalizada estación 21047010.....	36
Figura 3-7. Correlaciones Q_{90} adimensional vs Atributos.....	38
Figura 3-8 Correlaciones Q_{35} adimensional vs Atributos.....	38
Figura 3-9 Correlaciones Q_1 adimensional vs Atributos.....	39
Figura 3-10. Ejemplo de variación percentil a percentil del coeficiente de correlación.....	39
Figura 3-11. Esquema ilustrativo <i>K means</i> , izquierda: dispersión sin agrupar, derecha: grupos conformados.....	41
Figura 3-12. Análisis de sensibilidad <i>K means</i> con datos no estandarizados.....	42
Figura 3-13. Análisis de sensibilidad <i>K means</i> con datos estandarizados.....	42
Figura 3-14. Valores de R para Q_{35} vs valor atributo Clúster 3.....	44
Figura 3-15. Valores de R para Q_{35} vs valor atributo Clúster 7.....	44
Figura 3-16. Valores de R para Q_{35} vs valor atributo Clúster 11.....	45
Figura 3-17. Esquema de las Regiones Hidrológicas de Colombia.....	47
Figura 3-18. Distancia euclidiana al centroide del clúster.....	48
Figura 3-19. Agrupación de cuencas según precipitación (izquierda) y valor de precipitación media de la cuenca (derecha).....	49
Figura 3-20. Agrupación de cuencas según pendiente (izquierda) y valor de pendiente media de la cuenca (derecha).....	50
Figura 3-21. Agrupación de cuencas según ETP (izquierda) y valor de ETP media de la cuenca (derecha).....	50
Figura 3-22. Agrupación de cuencas según % de bosque (izquierda) y valor de % bosque de la cuenca (derecha).....	51
Figura 4-1. Variación del R^2 promedio en los diferentes percentiles de la CDC.....	60
Figura 4-2. Frecuencia de cada atributo en regresiones de dos variables.....	61
Figura 4-3. Variación de R^2 según percentil en regresiones con 2 o más variables.....	63
Figura 4-4. Frecuencia de cada atributo en regresiones de dos o más variables.....	63
Figura 4-5. Ejemplo de CDC real y estimada, estación 35017070 – Clúster 3, río Orotoy, área de drenaje: 167 km ²	64
Figura 4-6. Caudal adimensional real vs Caudal adimensional estimado por <i>RLM</i> en estación 35017070 y recta $y=x$	64
Figura 4-7. Ejemplo de CDC real y estimada, estación 42067010 – Clúster 9, río Vaupés, área de drenaje: 17077 km ²	65
Figura 4-8. Caudal adimensional real vs Caudal adimensional estimado mediante <i>RLM</i> en estación 42067010 y recta $y=x$	65
Figura 4-9. Ejemplo de CDC real y estimada, estación 21147080 – sub región Magdalena Cauca, río Cbrera, área de drenaje: 1185 km ²	66

Figura 4-10. Caudal adimensional real vs Caudal adimensional estimado mediante <i>RLM</i> en la estación 21147080 y recta $y=x$.	66
Figura 4-11. Comparación del error relativo entre la estimación mediante regionalización por clúster y por sub regiones.	69
Figura 5-1. Distribución espacial de estaciones de caudal con más de 30 años netos de registros.	76
Figura 5-2. Ejemplo de serie temporal descartada por ubicarse en lecho de múltiples canales.	77
Figura 5-3. Ejemplo de serie descartada por inconsistencias en las mediciones.	77
Figura 5-4. Distribución espacial del valor de la prueba H en estaciones de caudal.	78
Figura 5-5. Estaciones disponibles para el análisis no estacionario de la CDC.	79
Figura 5-6. Rengos de valores de pendiente de la recta de ajuste en estaciones analizadas.	80
Figura 5-7. Dispersión pendiente de Sen vs pendiente recta de ajuste.	81
Figura 5-8. Evolución temporal de la CDC y la varianza en el río Atrato.	82
Figura 5-9. Serie temporal Río Atrato entre 1976 y 2013, ajuste lineal $m=0.043 \text{ m}^3/\text{s}/\text{día}$.	82
Figura 5-10. Evolución temporal de la CDC y la varianza en el río Tubú.	83
Figura 5-11. Serie temporal Río Tibú entre 1973 y 2012, ajuste lineal $m=-0.0005 \text{ m}^3/\text{s}/\text{día}$.	83
Figura 5-12. Evolución temporal de la CDC y la varianza en el río Sogamoso.	84
Figura 5-13. Serie temporal Río Sogamoso entre 1960 y 2009, ajuste lineal $m=-0.0048 \text{ m}^3/\text{s}/\text{día}$.	84
Figura 5-14. Evolución temporal de la CDC y la varianza en el río Cauca.	85
Figura 5-15. Serie temporal río Cauca entre 1974 y 2013, ajuste lineal $m=0.0475 \text{ m}^3/\text{s}/\text{día}$.	85
Figura 5-16. Evolución temporal de la CDC y la varianza en el río Meta.	86
Figura 5-17. Serie temporal río Meta entre 1977 y 2013, ajuste lineal $m=0.006 \text{ m}^3/\text{s}/\text{día}$.	86
Figura 5-18. Evolución temporal de la CDC y la varianza en el río Manacacías.	87
Figura 5-19. Serie temporal río Manacacías entre 1978 y 2013, ajuste lineal $m=0.0064 \text{ m}^3/\text{s}/\text{día}$.	87
Figura 5-20. Evolución temporal de la CDC y la varianza en el río Putumayo.	88
Figura 5-21. Serie temporal río Putumayo entre 1971 y 2010, ajuste lineal $m=-0.0006 \text{ m}^3/\text{s}/\text{día}$.	88
Figura 5-22. Evolución temporal de la CDC y la varianza en el río San Pedro.	89

Figura 5-23. Serie temporal río San Pedro entre 1973 y 2013, ajuste lineal $m=0.002$ $\text{m}^3/\text{s}/\text{día}$	89
Figura 5-24. Evolución temporal de la CDC y la varianza en el río Patía.....	90
Figura 5-25. Serie temporal río Patía entre 1966 y 2004, ajuste lineal $m= -0.005$ $\text{m}^3/\text{s}/\text{día}$	90
Figura 5-26 Evolución temporal de la CDC y la varianza en el río San Juan.....	91
Figura 5-27. Serie temporal río San Juan entre 1973 y 2012, ajuste lineal $m=0.0283$ $\text{m}^3/\text{s}/\text{día}$	91

Lista de tablas

	Pág.
Tabla 2-1. Resumen de atributos.	30
Tabla 3-1. Atributos en cada tramo de la CDC.....	40
Tabla 3-2. Resumen de agrupación de cuencas calibradas	43
Tabla 3-3. Resumen de agrupación de cuencas sin calibrar.....	46
Tabla 4-1. Estaciones de validación para los clúster	55
Tabla 4-2. Estaciones de validación para sub regiones.....	55
Tabla 4-3. Resultados de <i>RLM</i> con 2 variables para cada clúster.....	57
Tabla 4-4. Resultados de <i>RLM</i> con 2 variables para cada sub región.....	58
Tabla 4-5. Ejemplo de resultados de <i>RLM</i> - clúster 8.....	58
Tabla 4-6. Ejemplo de resultados de <i>RLM</i> - sub región Magdalena Cauca.....	59
Tabla 4-7. Resultados generales de <i>RLM</i> en los clúster con 2 o más variables.	61
Tabla 4-8. Ejemplo de resultados de <i>RLM</i> clúster N° 8 con 4 variables.	62
Tabla 4-9. Valores de R y R^2 para observados - estimados vs $y=x$ en cada clúster.....	67
Tabla 4-10. Valores de R y R^2 para estimados - observados vs $y=x$ en cada sub región. .	67
Tabla 4-11. Errores relativos para validaciones en los clúster.....	68
Tabla 4-12. Errores relativos para validaciones en sub regiones.....	68
Tabla 5-1. Resumen de resultados del análisis no estacionario.....	92

1 Introducción

1.1 Generalidades

Para conocer el comportamiento del régimen hidrológico en lugares donde aún no se implementan mediciones o no se tienen registros suficientes, se necesita de una metodología que permita extrapolar o trasladar las condiciones desde los sitios donde sí se cuenta con información hacia aquellos desprovistos de datos. La modelación hidrológica se realiza principalmente por las limitantes asociadas a las técnicas de medición. No es posible aún calibrar todas las corrientes en cada uno de sus tramos, de hecho, aún se tiene un rango espacial y temporal de calibraciones muy reducido (Beven, 2012). En Colombia se han realizado esfuerzos en las últimas décadas por impulsar una red nacional de medición y toma de datos climáticos e hidrológicos con el fin de calibrar, estudiar y conocer el comportamiento del recurso agua en el país, sin embargo, aún permanecen muchas corrientes sin ningún tipo de registro o con muy escasas mediciones (Instituto de Hidrología Meteorología y Estudios Ambientales IDEAM, 2014).

Se hace necesario avanzar en el entendimiento de los procesos asociados al comportamiento de los caudales dentro de una cuenca hidrográfica y a la respuesta de la misma ante eventos de precipitación. Se conoce que las características y factores fisiológicos de cuenca como la topografía, el uso del suelo, la morfometría, la cobertura, el régimen climático entre otros, intervienen en el proceso de lluvia – escorrentía, además pueden servir como base para el entendimiento y la estimación de las descargas.

Para sintetizar y resumir el régimen de caudales a largo plazo de una corriente natural se utiliza uno de los conceptos en hidrología conocido como la Curva de Duración de Caudales o CDC, la cual representa la relación entre la magnitud y la frecuencia o el porcentaje de tiempo (de allí la palabra duración) que dicha magnitud es igualada o excedida, dada una corriente y un punto determinado. Ésta normalmente se trabaja en escala diaria, pero puede ser reconstruida con series de caudales en cualquier resolución temporal. Gráficamente, en el eje de las ordenadas se ubica la magnitud del caudal y en el eje de las abscisas el porcentaje de tiempo que el dato es igualado o excedido (Vogel & Fennessey,

1994). Debido a la cantidad de información que puede inferirse de ella y a los análisis del régimen de descargas que ésta posibilita, la CDC es uno de los conceptos y herramientas gráficas más utilizadas dentro de la hidrología junto con la curva de masas y la hidrógrafa (Foster, 1933).

La CDC se puede construir de forma empírica a partir de series de datos de caudales existentes en corrientes calibradas o instrumentadas, además de presentar la magnitud de las descargas en volumen por unidad de tiempo, ésta puede expresarse también en términos de caudal normalizado según el área de drenaje $[m^3/s/km^2]$ o en términos de caudal adimensional $Q^* = Q/\bar{Q}$ [adimensional] siendo \bar{Q} el caudal medio a largo plazo de la serie (Blöschl, 2013).

En el mundo, la mayoría de las cuencas hidrológicas se encuentran sin calibrar, a pesar de que para el año de 1995 ya habían más de 60.000 estaciones de medición de caudal en funcionamiento según la Organización Meteorológica Mundial WMS (WMO, 1995), se estima que alrededor de seis millones de cuencas permanecían para ese año sin calibrar o con información hidrológica insuficiente, lo que indica que solo alrededor del 1% de las cuencas se encuentran instrumentadas. Esto genera la necesidad de la utilización de modelos de estimación de series hidrológicas, o bien, del régimen de caudales en los sitios donde no se tienen dichos registros (Blöschl, 2005). Particularmente en el caso de Colombia, existe un enorme déficit en la cobertura de la red nacional de medición hidrológica, presentándose zonas y regiones totalmente desprovistas de información (Mesa Sánchez et al., 2003), por lo que se hace necesario desarrollar formas de estimación de CDC sintéticas con aplicación a nivel nacional. Si bien en otras regiones o países se tienen experiencias de regionalización de la CDC en el caso colombiano aún falta mucho por desarrollar y profundizar en este tema.

En este trabajo se aplica la regionalización de caudales a resolución temporal diaria, esta es la escala más utilizada en la reconstrucción de curvas de duración, sin embargo, ésta subestima el valor pico de caudal máximo instantáneo que se presenta en eventos de crecientes. Los eventos de máximos, se estima que se ubican en los percentiles inferiores a 2% de porcentaje de excedencia y particularmente, en cuencas que no superan los 500 km² de área de drenaje (Cataño-Álvarez & Vélez Upegui, 2016). La mayoría de unidades hidrológicas que aquí se trabajan sobrepasan dicho tamaño, por lo tanto se reduce considerablemente el número de corrientes a las cuales se les subestiman sus caudales de máximos, también se tiene que la mayoría de magnitudes estimadas, se encuentran en frecuencias mayores al 2% de probabilidad de excedencia (percentil 2), por lo que se tiene

un alto grado de confiabilidad en que la escala diaria represente bien los caudales ubicados entre los percentiles 2 y 100 (es decir, el 98% de los valores en una serie). Se conserva entonces la resolución temporal original, pues a pesar de la subestimación del caudal pico o $Q_{m\acute{a}x}$ instantáneo sigue teniendo gran utilidad para los medios científico o investigativo, ambiental e ingenieril, así como en el marco del ordenamiento territorial y la planeación y manejo del recurso hídrico, las captaciones y el diseño de plantas de tratamiento, las regulaciones energéticas y las operaciones en centrales hidroeléctricas, la navegación, entre otros.

Para compensar la escasez de información primaria (producto de aforos y mediciones) puede aplicarse una regionalización, la cual consiste en la división en sub partes o sub zonas del área de estudio, dichas zonas se escogen de tal manera que, las unidades hidrológicas (o sub cuencas hidrográficas) presenten similitud en aspectos o características referentes al comportamiento de los caudales (Mesa Sánchez et al., 2003). La regionalización de parámetros hidrológicos, puede definirse como la transferencia de información desde un sitio (cuenca hidrológica en este caso) calibrado o con información suficiente a otro que se encuentre sin calibrar, siempre y cuando se cumpla con la condición de que ambos se ubiquen dentro de una misma región hidrológica previamente definida (Razavi & Coulibaly, 2013). Para que dos territorios sean hidrológicamente homogéneos debe existir un alto grado de similitud en aspectos fisiológicos tales como clima, geología, topografía, vegetación, tipo de suelo predominante, entre otros. Puede decirse, que la clave del éxito en una regionalización de parámetros comienza por una rigurosa definición de cuáles cuencas son hidrológicamente homogéneas (Razavi & Coulibaly, 2013).

Es posible establecer un conjunto regional de parámetros o atributos correspondientes a las condiciones climáticas, topográficas y de paisaje que pueden describir en mayor o menor medida el comportamiento hidrológico de una cuenca (P. L. García, Méndez, & Zárate, 2017), existe una estrecha relación entre las características físicas de una unidad hidrológica y las variables que describen su respuesta lluvia – escorrentía, pues muchas de éstas se deben a la acción misma del agua (Mesa Sánchez et al., 2003). Se han estudiado los efectos de las características de una cuenca en la forma de su CDC, por ejemplo, se ha revisado si la geología y el clima de una determinada cuenca influyen de manera directa en el comportamiento de dicha curva (Musiake, Inokuti, & Talahasi, 1975) o si son relevantes los efectos del tipo de suelo predominante presentes en la cuenca (Eeles, Robinson, & Ward, 1990), o si tanto la topografía como el relieve son determinantes para la regulación hídrica y el régimen de descargas (Fennessey & Vogel, 1990). También se ha revisado si los cambios

en los tipos de coberturas vegetales influyen en el comportamiento de los caudales y en la respuesta de una determinada cuenca ante los eventos de precipitación (Burt & Swank, 1992), sin embargo, aún permanece como un reto para la comunidad hidrológica la profundización y el entendimiento de los fenómenos físicos asociados a cada uno de estos procesos y la relación de cada uno con la hidrología (Mohamoud, 2008).

Partiendo del hecho de que la variable caudal Q puede relacionarse directamente con los factores o atributos climáticos o de paisaje de una determinada unidad hidrológica, pueden ser estimados mediante regresión, los caudales correspondientes a cada uno de los percentiles característicos de la curva de duración y cuyo resultado estaría sujeto al error asociado a la dispersión de las relaciones entre los datos. La CDC también puede ser dividida en tres tramos, cada uno gobernado por procesos, fenómenos y variables que se asumen diferentes, siendo estas: la parte de la CDC que indica los caudales máximos que se ubican en los percentiles menores a 2% de probabilidad de excedencia está relacionada con la respuesta de la cuenca en forma de escorrentía superficial ante eventos extremos de precipitaciones de altas intensidades (crecientes, inundaciones, avenidas etc.), la parte de ésta que corresponde a los caudales de condiciones medias (entre 2% y 80% de probabilidad de excedencia) relata la competencia entre períodos secos y períodos de lluvia y por último, la parte relacionada con los caudales mínimos (entre 85% y 100% de probabilidad de excedencia) es gobernada por el flujo base o de recesión en períodos secos y está más influenciada por los procesos de recarga (Blöschl, 2013).

Se plantea la estimación de diferentes percentiles o caudales característicos Q_p de la CDC a partir, primero de la definición de las regiones hidrológicamente homogéneas del área de estudio definida como el territorio colombiano (límites políticos y administrativos del país) y la estimación mediante técnicas estadísticas de regresión. Se pueden aproximar los valores de los caudales dejando éstos en función de uno o varios atributos climáticos y de paisaje, suponiendo que:

$$Q_p = f\{X_1, X_2, \dots, X_n\} \quad (1-1)$$

siendo Q_p el caudal característico para un valor de porcentaje de excedencia p ($0 < p \leq 100$), y X_i el valor de un atributo - característica climática o de paisaje que le corresponde a la cuenca objeto de la estimación. Se propone entonces la estimación mediante regresión, de los parámetros que permitan calcular curvas de duración regionales en cada una de las zonas hidrológicamente homogéneas de Colombia.

En este trabajo, inicialmente se plantea el régimen de caudales como un proceso estacionario, es decir que sus condiciones y parámetros estadísticos son constantes a lo largo del tiempo. Tampoco es de interés el análisis en general de la variabilidad temporal (ciclos intra e inter anuales) sino que se asumen tanto las CDC reales como las estimadas como una representación del régimen de caudales a largo plazo. Sin embargo el tema de la no estacionariedad a largo plazo del comportamiento de los caudales en Colombia es tratado como capítulo aparte.

El clima es uno de los principales factores que influyen de manera directa en el régimen de descargas y en Colombia existen evidencias de los efectos del cambio climático en los valores de precipitación en varias regiones del país (Pérez, Poveda, Mesa, & Carvajal, 1998), por ende, es un tema pendiente de estudio, la estacionariedad o no del comportamiento a largo plazo de los caudales a través del tiempo. Hay experiencias a nivel global en el estudio de la estacionariedad de los caudales, se ha encontrado que en condiciones medias y mínimas no se han observado tendencias lo suficientemente claras, mientras que se tienen registros de que las inundaciones se han hecho más frecuentes (Barriendos et al., 2003; Milly et al., 2008; Villarini, Serinaldi, Smith, & Krajewski, 2009). Se espera entonces encontrar en este estudio, principalmente variaciones en los percentiles correspondientes a los caudales máximos (menores a 2% de probabilidad de excedencia) en la CDC.

Los cambios temporales en la CDC pueden ser observados mediante un análisis de largo plazo, el análisis de tendencias planteado difiere de un análisis estacional, es decir, el que se enfoca en los ciclos temporales como el intra e inter anual. El problema de la estacionalidad dificulta el análisis de tendencias pues induce variaciones en los valores de las series, en este caso, de caudales (Gilbert, 1987), por lo que se propone un ajuste lineal a series con suficiente longitud temporal de registros (más de 30 años) y un análisis de tendencias mediante la prueba no paramétrica de Mann – Kendall (Gilbert, 1987; Mann, 1945).

1.2 Justificación

La predicción de caudales en cuencas con información escasa es una de las tareas más importantes para el estudio, manejo y planeación del recurso hídrico a nivel global, ésta permanece aún hoy, como uno de los retos más importantes para la comunidad hidrológica (Sivapalan, 2003).

En el caso de Colombia, sería de gran utilidad contar con estimaciones del comportamiento del régimen de caudales en las cuencas o territorios que se encuentran desprovistos de

información, representando este mediante la curva de duración de caudales diaria. Existen numerosas aplicaciones en ingeniería y planeación ambiental, territorial para dichas CDC, como lo son por ejemplo, el análisis de caudales mínimos y máximos para diferentes períodos de retorno, la estimación de caudales ambientales o ecológicos para el cálculo de la oferta hídrica en estructuras de captación, la estimación y zonificación de la oferta hídrica en los planes de ordenamiento y manejo del recurso hídrico POMCA y PORH. Además siendo Colombia un país con un gran potencial hidroenergético, el conocimiento acerca del régimen de descargas puede ser de gran utilidad para el diseño, la ejecución y la operación de pequeñas y grandes centrales hidroeléctricas proyectadas y existentes en las cuencas hidrográficas sin los registros suficientes como para ser consideradas como calibradas pues la CDC representa una herramienta de gran utilidad cuando se trata de intervenciones antrópicas en los ríos.

En este trabajo se enfoca en la realización de diferentes aportes a la disciplina de la hidrología, éstos constan no solo del desarrollo de una herramienta útil, sino también en la contribución al entendimiento de las relaciones existentes entre el régimen de caudales de una cuenca y diferentes características morfométricas y variables asociadas al clima, la fisiología y el paisaje. También se plantea realizar un aporte enfocado en el análisis y la delimitación de las regiones hidrológicamente homogéneas en Colombia y cómo esto puede contribuir a robustecer la estimación del régimen a largo plazo.

Muchos sistemas de planeación del agua han sido diseñados bajo el principio de la estacionariedad, éste ha sido asumido en diferentes aplicaciones hidrológicas y distintos enfoques tales como frecuencia y magnitud de las inundaciones, oferta hídrica, calidad del agua, entre otros. La estacionariedad de parámetros climáticos debido a cambios antrópicos y a ciclos geoclimáticos (IPCC, 2014), genera la necesidad de estudiar el ciclo del agua, pero teniendo en cuenta las posibles variaciones temporales, a largo plazo, en algunos parámetros relacionados con los procesos de lluvia y escorrentía, en especial aquellas llamadas variables climáticas. El entendimiento de la respuesta de algunas cuencas del país ante posibles cambios graduales a lo largo del tiempo se hace posible con un análisis no estacionario, el cual puede ser llevado a cabo mediante la revisión de variaciones temporales en las curvas de duración de diferentes ríos.

1.3 Antecedentes

1.3.1 Definición de regiones hidrológicas

Se ha planteado la agrupación de cuencas hidrológicas con el fin de obtener mejores resultados de aproximación en cálculos hidrológicos. En la región de Mixteca Oaxaqueña, al sur de México (Álvarez, Hotait, & Sustaita, 2011), agruparon 17 subcuencas a partir de 20 variables o atributos tanto climáticos como morfométricos para delimitar regiones hidrológicamente homogéneas. El análisis tuvo como fin, estimar la oferta media superficial en las diferentes unidades hidrológicas no calibradas de la región.

Sawicz, Wagener, Sivapalan, Torch, & Carrillo, (2011) estudiaron como clasificar un grupo de cuencas según similitud de diferentes características referentes a su respuesta ante eventos de precipitación. Para ello, utilizaron información de 280 corrientes instrumentadas, ubicadas en la región este de los Estados Unidos. A éstas se les observó relación entre el régimen de caudales (curva de duración) con cuatro índices: relación de lluvia, índice de flujo base, relación de días nevados, pendiente de la CDC, elasticidad del flujo, y elasticidad creciente de las ramas. Mediante un análisis de clúster jerárquico se agruparon las unidades seleccionadas según cada uno de los índices. El set de resultados arrojó que existe una alta heterogeneidad hidrológica en la región de estudio, generalmente, las cuencas pertenecientes a un mismo grupo, no presentaron vecindad entre sí.

Sauquet & Catalogne, (2011) realizaron un resumen de metodologías de agrupación de cuencas con el fin de obtener mejores resultados a la hora de estimar la CDC en sitios sin información, con este fin, seleccionaron aproximadamente 3500 cuencas con información de caudales diarios disponibles ubicadas en el territorio continental de Francia y en la isla de Córcega, todas ellas con por lo menos 18 años de registro. En el trabajo se señala, que la correcta agrupación de cuencas permite obtener mejores resultados de regionalización, pero que además, no existe una metodología universal para delimitar estas regiones. Se utilizó el análisis de componentes principales *PCA* con el fin de sintetizar la información proveniente de los atributos. Se encontró que la similitud hidrológica puede asociarse a dos indicadores simples, el índice de conectividad (*IC*), el cual representa la forma de la CDC y depende de tres caudales característicos Q_p y del ratio de estacionalidad (*SR*) que representa la relación entre caudales en períodos secos y caudales en períodos de lluvia. Los resultados mostraron que los modelos de agrupación basados en regresión de parámetros y atributos, generalmente brindan buenos resultados de estimación del régimen de caudales.

Waseem, Ajmal, & Kim, (2016) propusieron una mejora a los resultados de aproximación de la CDC mediante una técnica de regresión, pero acompañada previamente de la definición de regiones hidrológicamente homogéneas, para ello, plantean relaciones entre los atributos de 8 cuencas hidrográficas vecinas ubicadas al noreste de Pakistán. Se utilizaron los métodos de la diferencia media aritmética (*DM*) y la media de discordancia basada en momentos *L* (*DL*) que cuantifican la similitud entre las características de cada cuenca. Según el test de homogeneidad aplicado a los atributos geomorfológicos y climáticos, las cuencas estudiadas presentaron un 77.4% de similitud, en promedio, para los grupos conformados.

P. L. García et al., (2017) realizaron una agrupación de cuencas o *clúster* con el fin de definir las regiones hidrológicamente homogéneas de Colombia, por lo que se definió un umbral de área y se dividió el territorio nacional en 9139 cuencas hidrográficas, a cada una se le atribuyeron 20 características climáticas, y topográficas. Utilizaron análisis de componentes principales o *PCA* con el fin de crear nuevos grupos de variables que explican las variables originales, los resultados arrojaron cinco componentes que dependían principalmente de variables climáticas (precipitación y evapotranspiración). Mediante juicio experto, definieron que el territorio colombiano podría dividirse en un total de 16 regiones hidrológicamente homogéneas y utilizando el algoritmo de *K-means clustering* realizaron la agrupación de las cuencas colombianas.

1.3.2 Estimación de curvas de duración regionales

La estimación del régimen de descargas ha tenido solución en diferentes locaciones geográficas y por diferentes autores en el mundo, diversos enfoques y metodologías han sido utilizados con el fin de llevar a cabo ésta aproximación. Por ejemplo, (Mesa Sánchez et al., 2003), aplicaron una metodología basada en regresiones potenciales simples, obtenidas a partir de variables geomorfológicas, para encontrar relaciones regionales de características medias y máximas para caudales normalizados según el área de drenaje de sus respectivas cuencas. La regionalización se realizó para las 5 sub regiones de Colombia. En general, los autores obtuvieron valores de errores porcentuales de entre 10 y 50% aproximadamente, comparando los resultados de la regionalización con los caudales observados. La estimación se realizó para crecientes con períodos de retorno de 2.33, 5, 10, 25, 50 y 100 años.

Krasovskaia, Gottscalk, Leblois, & Pacheco, (2006) regionalizaron CDC diarias en Costa Rica mediante un modelo de estimación descrito según el valor medio de un coeficiente de variación y la comparación de los datos observados con los simulados. La estimación de los

errores arrojó que éstos oscilan entre aproximadamente 10 y 30%, correspondiendo los valores más altos a los caudales mínimos y máximos (percentiles mayores a 85% y menores a 20% en probabilidad de excedencia) y las incertidumbres más bajas para las condiciones medias.

Se propuso un problema de regionalización del modelo *Index - Flow* a partir de 18 CDC reales con resolución temporal diaria, correspondientes a sendas corrientes calibradas en la región central de Italia (Castellarin, Camorani, & Brath, 2007). El modelo fue utilizado para calcular tanto CDC anuales como de largo plazo. Las mejores aproximaciones en este caso se obtuvieron para los caudales medios y máximos, mientras que las estimaciones de caudales mínimos presentaron los más altos valores de incertidumbre.

Mohamoud, (2008) regionalizó las curvas de duración de caudales diarias para cuatro regiones hidrológicas previamente definidas localizadas en la zona este de Norteamérica a partir de la reconstrucción de curvas de duración de caudales normalizados según el área de drenaje. La estimación se realizó mediante un modelo regional basado en la regresión lineal múltiple o *RLM*, cuyas variables independientes fueron obtenidas a partir de una serie de atributos. El autor propone la reconstrucción de la CDC a partir de la estimación de 15 caudales característicos Q_p producto de la validación de ecuaciones compuestas por un coeficiente y dos exponentes obtenidos mediante la *RLM*. Fueron obtenidos buenos resultados de aproximación, con valores del coeficiente R^2 en las regresiones múltiples de entre 0.74 y 0.99.

Li, Shao, Zhang, & Chiew, (2010) evaluaron el desempeño del modelo semiparamétrico *Index*, el cual se basa en el establecimiento de relaciones no paramétricas entre las variables que intervienen en los procesos asociados a la escorrentía, para posteriormente aplicar regresiones lineales y realizar la estimación de los caudales a partir de éstas. El método fue aplicado a 227 cuencas ubicadas en el este de Australia y arrojó resultados satisfactorios con valores de coeficientes de correlación generalmente mayores a 0.8, al igual que en otros trabajos revisados, los mejores resultados se obtuvieron en las aproximaciones de caudales en régimen medio.

Castellarin, (2014), planteó abordar la estimación de la curva de duración en cuencas sin información mediante una interpolación empírica tridimensional. Estando las dos primeras variables (x, y) dadas en función de once atributos de la cuenca, y la tercera variable (z) en función de diferentes percentiles característicos normalizados. La interpolación fue realizada

para 41 cuencas sin calibrar distribuidas en las regiones centro y norte de Italia. Se obtuvieron valores del coeficiente de Nash bastante aceptables (entre 0.85 y 0.95).

Gallego & Carvajal, (2015) realizaron una regionalización para diferentes tramos de la curva de duración de caudales en seis regiones hidrológicas previamente definidas por la Unidad de Planeación Minero Energética (UPME-PPUJ, 2015) para el departamento de Antioquia (noroeste de Colombia). La estimación de los diferentes caudales de la CDC se realizó mediante una relación lineal entre el valor de Q en cada percentil característico con el caudal medio y la probabilidad de excedencia o la frecuencia. En general el error relativo obtenido fue inferior al 13% para todas las regiones con excepción de la región del Urabá cuyo error porcentual relativo fue de 44%.

Salazar Oliveros, (2016) realizó tanto la agrupación de cuencas como la estimación de las CDC regionales en la región de Santander y Norte de Santander en el noreste de Colombia, se seleccionaron 39 ríos que contara cada uno con estación de la red de monitoreo del IDEAM, para agruparlos según similitud en atributos mediante el algoritmo *k - means clustering* y las Curvas de Andrews. En cada uno de los 8 grupos definidos, se estimó la CDC regional a partir de dos tipos de regresión, lineal y no lineal, obteniendo considerablemente mejores resultados en las regresiones no lineales. En general los valores del coeficiente de Nash obtenidos por el autor eran en su mayoría cercanos a 1, esto indica que la estimación del régimen, habiendo definido previamente las regiones hidrológicas, es coherente y puede arrojar resultados satisfactorios.

Se han desarrollado otras metodologías para la predicción del régimen de caudales en sitios sin información, por ejemplo, (Atieh, Taylor, Sattar, & Gharabaghi, 2017) presentaron predicciones realizadas a partir de Redes Neuronales Artificiales (*ANN*) y Programación de Expresión Génica (*GEP*) utilizando 171 cuencas sin calibrar y 89 cuencas calibradas ubicadas en Norteamérica. Se utilizaron las características climáticas, fisiológicas y topográficas con el fin de identificar relaciones entre éstas y parámetros relacionados con el régimen de caudales (media y varianza). Los resultados mostraron valores del coeficiente de correlación de entre 0.9 y 0.95 comparando las estimaciones con los datos de caudal (caudales observados).

1.3.3 Análisis no estacionario

Existen numerosas evidencias de que las condiciones climáticas evolucionan de forma trascendente a los ciclos establecidos, es decir que existen tendencias a largo plazo en

variables relacionadas con el ciclo hidrológico. (Pérez et al., 1998) implementaron técnicas estadísticas para identificar señales de cambio climático mediante el análisis de series de datos hidrológicos en Colombia. Analizando las variables temperatura, precipitación y caudal, encontraron que existe una clara tendencia al calentamiento en los valores medios y mínimos de temperatura (tal y como ocurre de manera generalizada en el resto del mundo). Las series de precipitación no presentaron un comportamiento uniforme, pues se encontraron tanto tendencias crecientes como decrecientes, mientras que en las series de caudales, se encontró que para las principales cuencas del país, existen leves tendencias decrecientes.

Para la cuenca media y alta del río Cauca, (Puertas, Carvajal, & Quintero, 2011) mediante análisis de tendencias y la prueba de Mann-Kendall, detectó una disminución en los valores mensuales de precipitación para algunas épocas del año, mientras que (M. C. García, Botero, Quiroga, & Robles, 2012) estudiaron la variabilidad y el cambio climático enfocado en el manejo del recurso hídrico colombiano, revelando una alta relación entre los fenómenos de variación climática a largo plazo y el uso del recurso agua en el país. Recomendaron tener en cuenta los análisis de no estacionariedad en las políticas ambientales y de planeación en Colombia.

Barriendos et al., (2003) analizaron series de registros de inundaciones en la península ibérica desde los siglos XIV y XVII hasta la fecha, en general se encontró que las inundaciones son más frecuentes y poseen mayor magnitud, además, ninguna de las cuencas estudiadas tanto en España como en Francia presentaron estacionariedad en el régimen de caudales. Por su parte, (Villarini et al., 2009) estudió la estacionariedad de los picos anuales de inundaciones en 50 estaciones de medición en ríos de Estados Unidos con al menos 100 años de registros durante el siglo 20. Se encontraron cambios significativos en los valores medios anuales para la mayoría de las estaciones. (Westra & Sisson, 2011) también detectaron no estacionariedad en eventos extremos de precipitación en Australia. En Reino Unido, (Prosdocimi, Kjeldsen, & Svensson, 2013) encontraron evidencias de cambio climático en series de precipitación, éstas presentaron tendencias crecientes tanto en el valor como en la frecuencia de eventos extremos.

Con referencia a cambios en la curva de duración de caudales, (Brown, Zhang, McMahon, Western, & Vertessy, 2005) presentaron un resumen de casos de estudio en Australia y Nueva Zelanda de los efectos provocados por cambios en las coberturas y los usos del suelo en el comportamiento de los caudales. Para este estudio se tuvieron en cuenta cuencas con registros suficientemente largos de caudales además de información correspondiente a datos

históricos de variables referentes a deforestación, reforestación, urbanización, entre otras. Los resultados señalan que de forma generalizada, se ha subestimado el potencial que tienen las variables asociadas a la cobertura de una cuenca para regular la respuesta hidrológica y controlar el régimen de caudales o la forma de la CDC. (Milly et al., 2008) proclaman la muerte de la estacionariedad asumida y aplicada en la mayoría de los sistemas de planeación del recurso hídrico en el mundo, afirman que la cantidad de cambios antrópicos que han tenido lugar en la tierra han hecho evolucionar, en el tiempo, variables como precipitación, caudales, inundaciones, niveles del mar, ente otras. Los modelos Lluvia – Escorrentía también han sido tópico de interés para el análisis no estacionario, (Vaze et al., 2010) utilizó los parámetros de cuatro modelos de este tipo calibrados con series históricas de caudales en cuencas australianas. Los resultados allí obtenidos, indican que es recomendable calibrar los modelos con series recientes, si se calibran con series mayores a 20 años, las respectivas medias no deben variar más de un 15%.

En Colombia se han estudiado los cambios a largo plazo de la variable precipitación, (Cantor & Ochoa, 2011) evaluaron tendencias en series de precipitación a escala mensual, días de lluvia mensuales y series mensuales de lluvia promedio diaria en 45 estaciones hidroclimatológicas ubicadas en Antioquia. Las tendencias de estas series fueron estimadas a partir de diferentes pruebas: (T de Sen, Mann-Kendall, Rho de Spearman) para posteriormente realizar la prueba de homogeneidad en la tendencia de Van Belle y Huges. Se observaron tendencias crecientes en la mayoría de estaciones para las cuales se realizó el análisis. Por otra parte, (Urán, 2016) realizó un estudio de la lluvia en Colombia a diferentes escalas temporales utilizando datos de la misión satelital TRMM (*Tropical Rainfall Measure Mission*) y series de lluvia provenientes de estaciones meteorológicas en tierra. Los análisis arrojaron que para análisis con resoluciones temporales mayores a doce horas de la variable precipitación, las tendencias no son significativas, mientras que sí lo son para escalas temporales menores.

Estupiñán, (2016), profundizó en el estudio de la variabilidad espacial y temporal de la precipitación en Colombia mediante el cálculo de modos de oscilación por medio de un análisis espectral singular y la transformada de *Hilbert*. El estudio se llevó a cabo para regiones hidrológicas obtenidas a partir del algoritmo de *k-means* agrupando según los valores medios de la precipitación, los datos de lluvia utilizados fueron de estaciones de medición en tierra del IDEAM y las misiones TRMM y GPCC. El análisis de tendencias mostró que la precipitación anual decrece en el 15% de las estaciones analizadas, aumenta en otro 15% y en el 70% restante, no se presentaron tendencias significativas.

1.4 Objetivos

1.4.1 Objetivo general

Construir curvas de duración de caudales a partir de una regionalización hidrológica en Colombia con aplicación a cuencas con información escasa.

1.4.2 Objetivos específicos

- Proponer una metodología adecuada para la definición de regiones hidrológicas en Colombia.
- Estimar el régimen de caudales para cada región hidrológica de Colombia.
- Estudiar el comportamiento temporal de la curva de duración de caudales en algunas cuencas del país.

1.5 Metodología

De manera general, la metodología planteada para la realización de este trabajo consta de cuatro fases principales, la primera trata de la recopilación, el manejo y procesamiento de la información, la segunda consiste en la definición de las regiones hidrológicas en Colombia para posteriormente utilizarlas como insumo previo a los cálculos de regresión enfocados a la estimación del régimen hidrológico en cada una. Por último, se revisa la estacionariedad de la CDC utilizando algunas de las estaciones hidrológicas con suficientes registros mediante un análisis de tendencias.

Se buscan y organizan series de caudales a escala diaria con las cuales es posible generar CDC reales, también se obtiene información topográfica climática, geológica y de cobertura y uso del suelo. Se trazan las cuencas correspondientes a las estaciones seleccionadas para el estudio, y a cada una se le asigna un set de atributos y descriptores clima – paisaje para posteriormente, con ellos, realizar un clúster de tipo *K-means*. Se divide también la totalidad del territorio nacional en sub cuencas o unidades hidrológicas con el fin de extrapolar los resultados del clúster y definir las regiones hidrológicamente homogéneas de Colombia.

Utilizando como insumo de calibración las CDC reales, la distribución de las cuencas en los clúster y los atributos a tener en cuenta en cada una, se estiman los valores para diferentes

caudales característicos con el fin de generar curvas de duración diarias, sintéticas y regionales para cada agrupación o región hidrológica, por lo que se propone una regresión lineal múltiple en su forma potencial:

$$Q_p = \beta X_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n} \quad (1-2)$$

donde Q_p corresponde a cada uno de los percentiles característicos de la CDC, X_i corresponde a los atributos climáticos, topográficos, geológicos de cobertura y morfométricos, β_i y α corresponden al coeficiente y los exponentes producto de regresión respectivamente. Los resultados de la regionalización son evaluados comparando los Q_p de las CDC observadas, no tenidas en cuenta en los cálculos con los estimados mediante el modelo.

El análisis no estacionario se realiza mediante la selección de series de caudales con más de treinta años netos de registro y se realiza la prueba no paramétrica de Mann Kendall para verificar la existencia de tendencias a largo plazo, en caso de haber, se calcula la pendiente de la regresión lineal de la serie de datos, además del estimador de pendiente de Sen. Se grafican posteriormente las CDC correspondientes a diferentes dominios temporales en los casos más llamativos, superpuestas en una misma figura. Se busca comparar las curvas de duración construidas con diferentes fragmentos temporales de una serie, por lo que ese escoge diferentes percentiles característicos, se calcula su valor anual y se crean series temporales con sus respectivas tendencias lineales.

Con el fin de complementar el análisis no estacionario, se cuantifica la tendencia en la varianza de las series seleccionadas para el análisis, para ello, calcula el valor de la desviación estándar en una ventana móvil de 10 años con paso anual, posteriormente se calcula la pendiente de la recta de ajuste de mínimos cuadrados. Se asume de entrada, que los cambios temporales en esta variable ocurren con tendencia lineal.

2 Datos

Para la realización de este trabajo es necesario contar con datos de diversas características, inicialmente, se requieren series de caudales suficientemente largas y a una escala temporal favorable para la elaboración de la CDC de largo plazo. Entre los diferentes formatos de información de caudales líquidos existentes para los ríos de Colombia, se elige la escala diaria pues se cuenta con suficientes registros tanto temporalmente como espacialmente y porque puede representar de forma adecuada la curva de duración. Habiendo realizado una revisión preliminar de cuáles pueden ser las variables a escala de cuenca que intervienen o están relacionadas con los proceso de lluvia – escorrentía de los cuales deriva el regímenes de caudales, se obtiene información hidrológica, cartográfica, información topográfica satelital, entre otras, como insumos necesarios. En este capítulo se listan los datos recolectados y su respectivo manejo y procesamiento.

2.1 Información

2.1.1 Caudales

La información de caudales líquidos a escala diaria es extraída de la base de datos de la red nacional de medición del Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia IDEAM, estos datos se encuentran libres para su descarga gratuita en la plataforma virtual SIRH (sirh.ideam.gov.co). Se obtiene una base de datos con 849 series de caudal líquido a escala diaria distribuidas espacialmente por todo el país. Entre todas las estaciones se cuenta registros desde el año 1940 hasta el año 2015. Las mediciones son discontinuas, es decir que existen períodos de tiempo sin registros durante la vida útil de las estaciones (presentan datos faltantes). La red Nacional de medición de caudales del IDEAM, tiene sus puntos de monitoreo distribuidos según se muestra en la Figura 2-1.

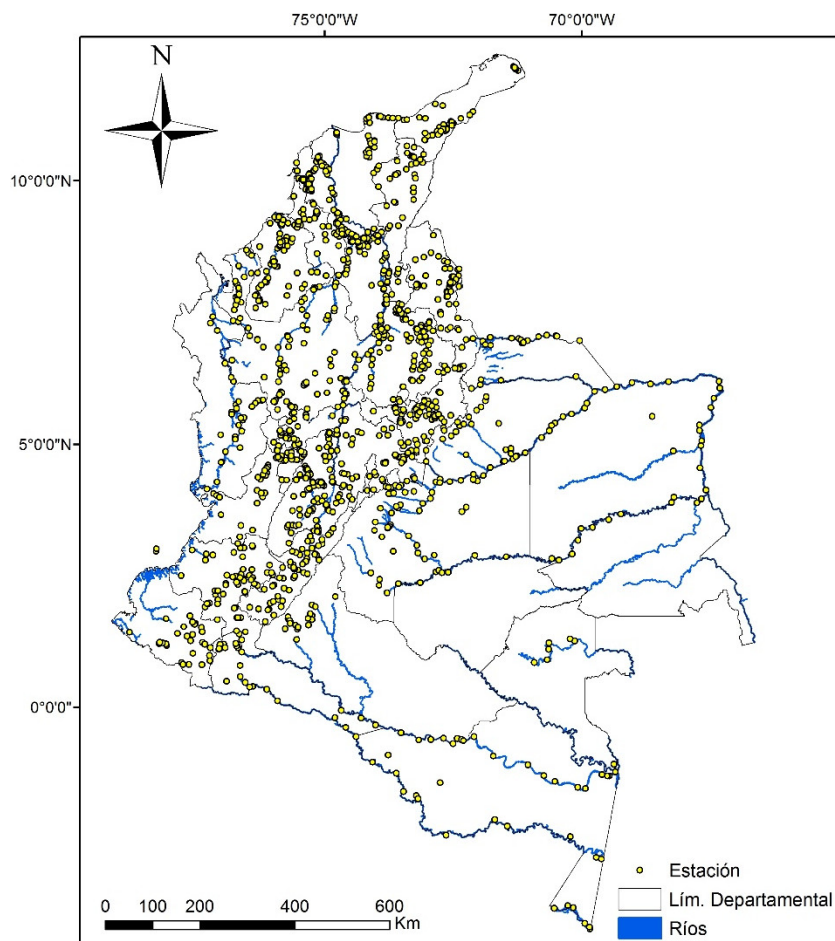


Figura 2-1. Estaciones de caudales diarios en Colombia, red nacional de mediciones del IDEAM.

2.1.2 Modelo digital de elevación

Se requiere de un modelo topográfico o de elevación digital de terreno (*DEM*) que abarque toda el área continental de Colombia, éste es el principal insumo para delimitar las cuencas y obtener los diversos parámetros y variables topográficas de cada una. Se recurre al modelo de *SRTM* 90 de la misión topográfica global de la Agencia Nacional de Inteligencia Geoespacial *NGA* y la Administración Nacional de la Aeronáutica y del Espacio, *NASA* iniciada en el año 2000 que levantó información de elevación en metros sobre el nivel del mar (msnm) con resolución de 90 metros de tamaño de pixel, esta información se encuentra disponible para latitudes entre los 56° sur y los 60° norte. Los datos se componen de diversas planchas descargables de forma gratuita desde el portal web: <http://srtm.csi.cgiar.org>. El tamaño de celda de 90 m por 90 m es ideal para la realización de los cálculos en términos

de gasto computacional, debido a que el dominio espacial planteado como zona de estudio de este trabajo es de aproximadamente $1 \times 10^6 \text{ km}^2$, por lo que se opta por esta resolución y no por otras de menor tamaño de pixel, las cuales también se encuentran disponibles para descargar de forma gratuita como lo son SRTM de 30 m y Vertex Alos – Palsar de 12.5 m. La Figura 2-2 muestra las planchas con información topográfica de Colombia unidas en un solo mosaico.

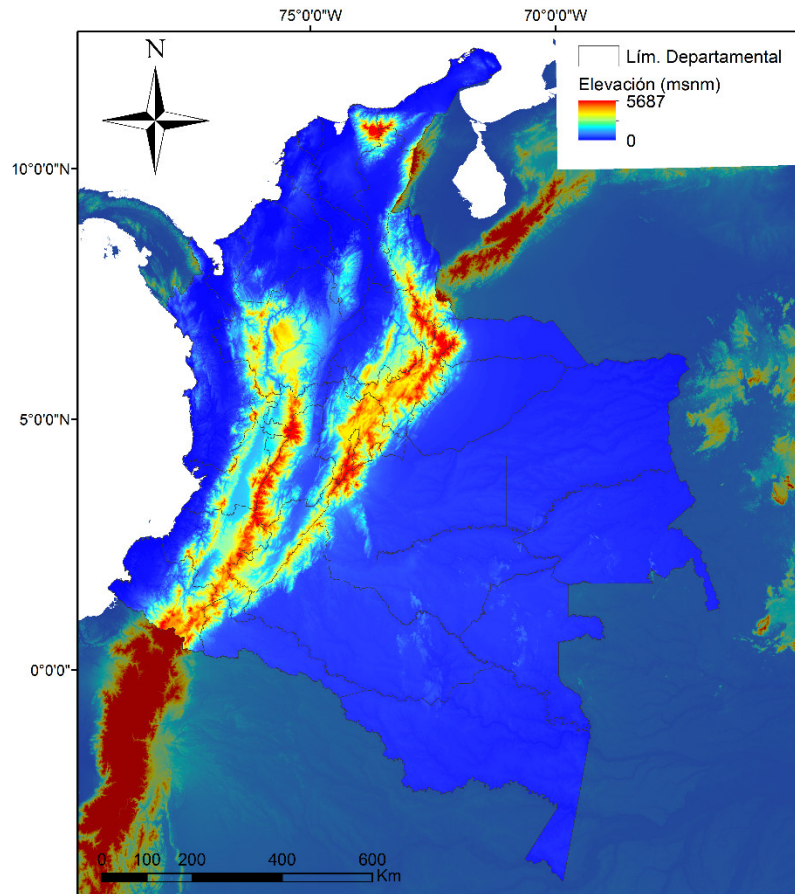


Figura 2-2. Modelo de Elevación Digital de Colombia - resolución 90 m, fuente: *SRTM*.

2.1.3 Mapa de coberturas

El Instituto Geográfico Agustín Codazzi *IGAC* a través del Sistema de información geográfica para el ordenamiento territorial *SIG-OT*, brinda información geográfica y cartográfica de todo el territorio colombiano, esta información está disponible para descargar desde la página <http://sigotn.igac.gov.co>. Las variables o atributos referentes al tipo de

cobertura del suelo, se encuentran disponibles en dicho portal en un mapa escala 1:100.000. En la Figura 2-3 se delimitan áreas agrícolas, áreas de bosque y áreas urbanas.

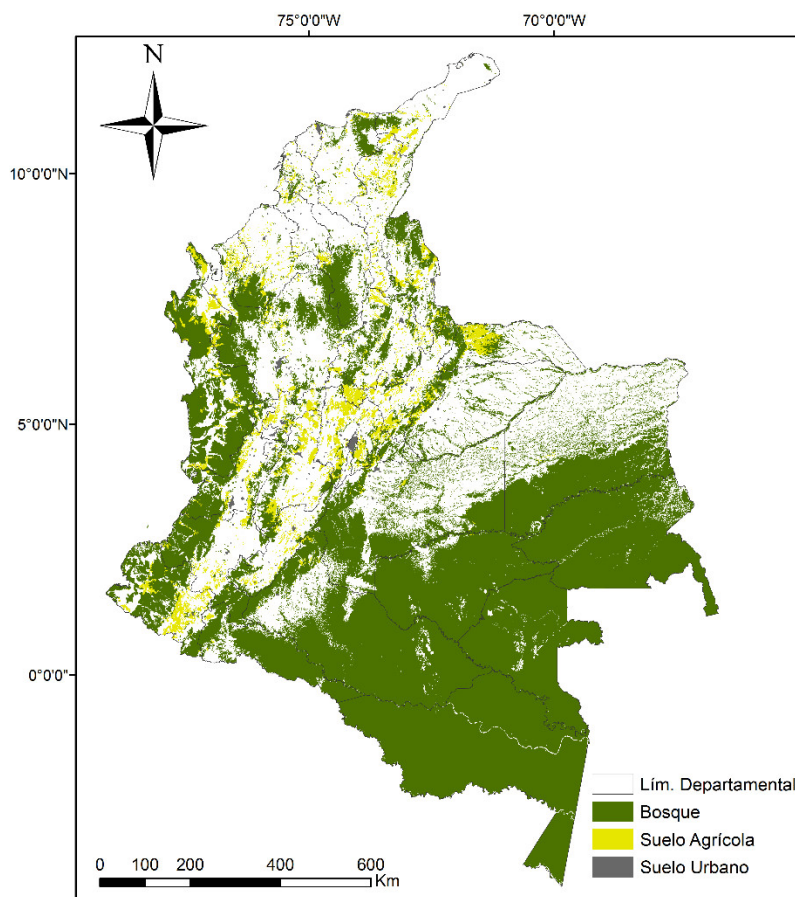


Figura 2-3. Mapa de coberturas de Colombia, escala: 1:100.000, fuente: *IGAC*.

2.1.4 Mapa geológico de Colombia

Para obtener los atributos geológicos de cada cuenca, particularmente del tipo de material predominante en el suelo, se utiliza el levantamiento geológico compilado en el 2014 por el Instituto Geográfico Agustín Codazzi *IGAC* y el Servicio Geológico Colombiano *SGC*, en la Figura 2-4 se presenta un esquema con las divisiones de los distintos tipos de suelo predominantes en el país, en este caso los más comunes utilizados en la bibliografía: suelos arenosos, limosos y arcillosos. Junto con la información de geología, de las mismas fuentes, se obtiene el mapa de fallas tectónicas de Colombia, también en escala 1:100.000 (Figura 2-5).

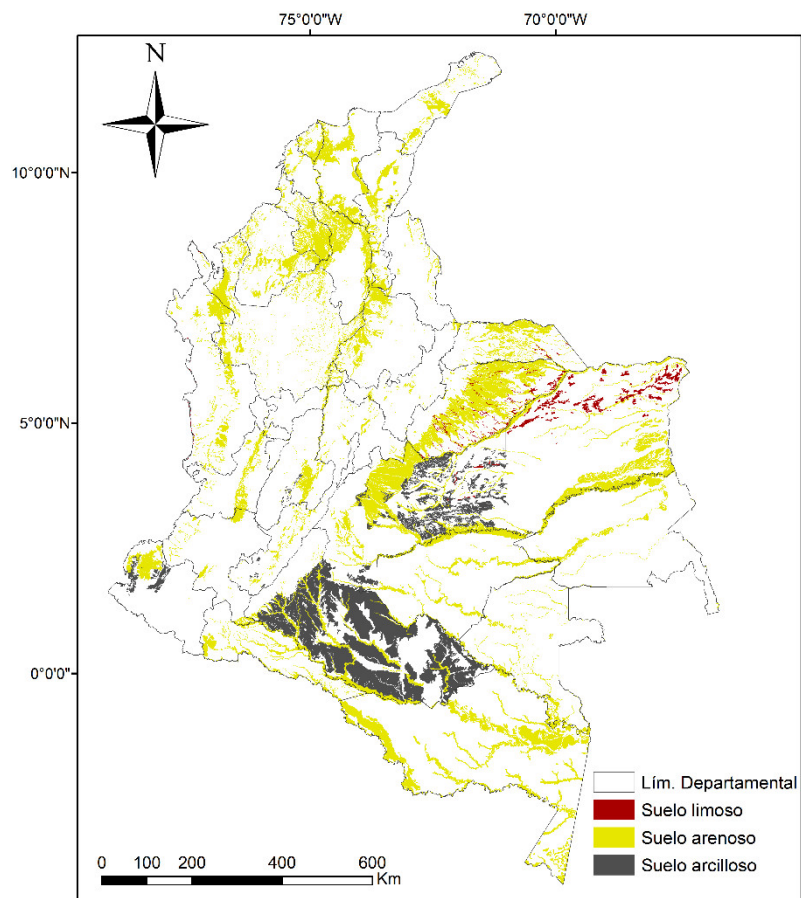


Figura 2-4. Mapa geológico de Colombia, escala 1:100.000, fuente: IGAC - SGC.

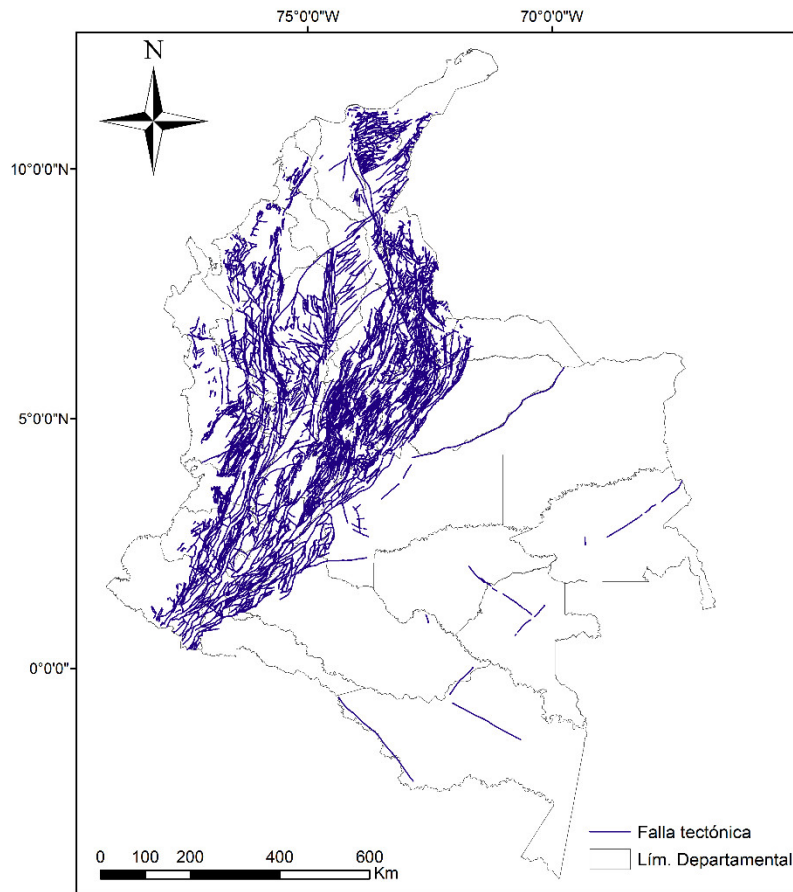


Figura 2-5. Mapa de fallas geológicas de Colombia. Escala: 1:100.000, fuente: *IGAC-SGC*.

2.1.5 Campos de precipitación mensual de Colombia

Hurtado & Mesa, (2014) presentaron la reconstrucción de los campos históricos de precipitación mensual en Colombia entre los años 1975 y 2006 a una resolución espacial de cinco minutos de arco (aproximadamente 9.3 km). Los mapas mensuales fueron estimados mediante la integración de datos de lluvia en tierra y datos satelitales. Se calcula el promedio de todos los meses comprendidos en este período (Figura 2-6) como valor de precipitación media mensual (mm/mes) y el valor máximo de precipitación en cada elemento espacial (Figura 2-7) como precipitación máxima mensual en mm/mes.

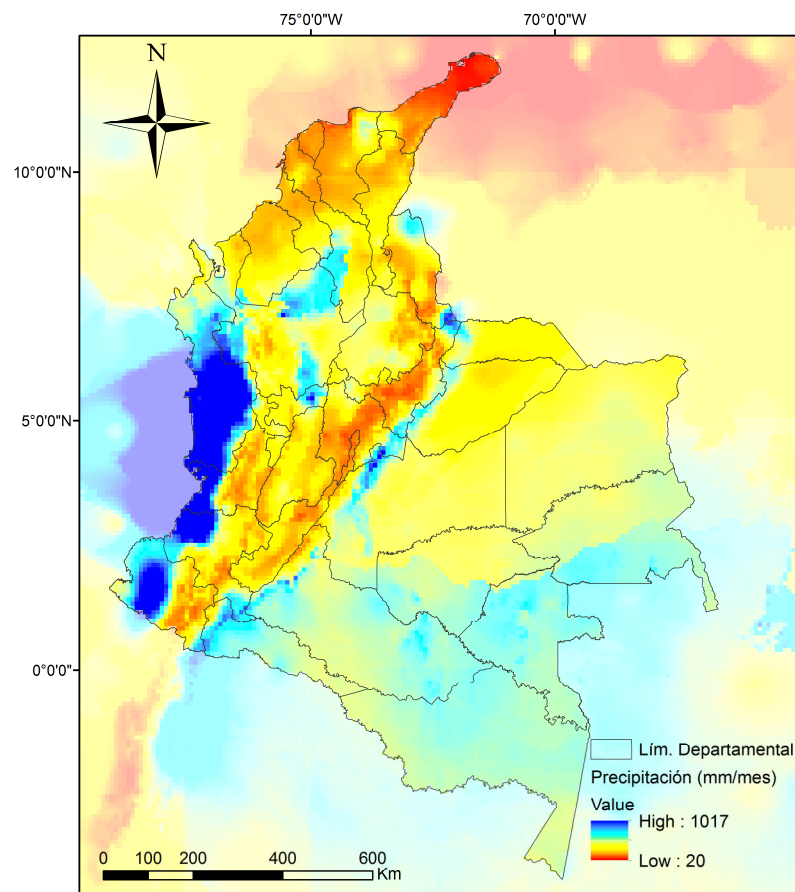


Figura 2-6. Precipitación media mensual de Colombia 1975 – 2006, fuente: Hurtado y Mesa (2014).

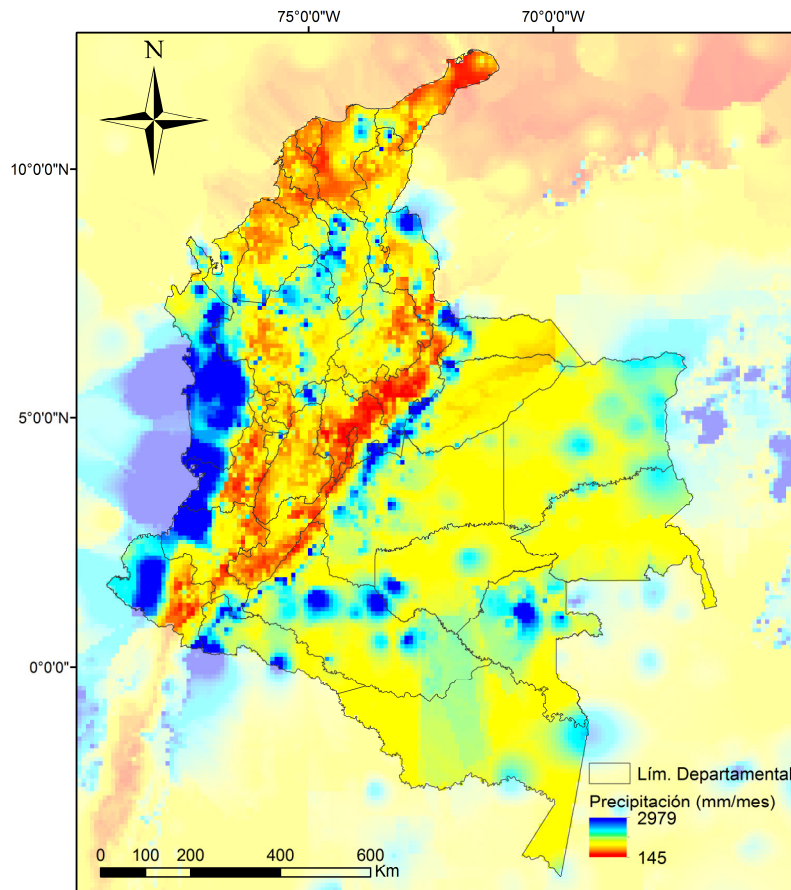


Figura 2-7. Precipitación máxima mensual de Colombia 1975 – 2006, fuente: Hurtado y Mesa (2014).

2.1.6 Evapotranspiración potencial

La estimación de la evapotranspiración potencial ETP se calcula mediante la ecuación de Cenicafé propuesta por (Jaramillo, 1989) en la cual la ETP anual depende únicamente de la elevación sobre el nivel del mar, la ecuación (2-1) es producto de una regresión entre valores de evaporación provenientes de estaciones climáticas y la elevación del terreno.

$$ETP = 1017.17 \exp(-0.0002 h) \quad (2-1)$$

Siendo ETP la evapotranspiración potencial media anual en milímetros por año y h la elevación en metros sobre el nivel del mar. A pesar de ser una estimación con incertidumbre asociada a la dispersión, se estima la variable ETP con este procedimiento debido a su

sencillez y a que se cuenta con información de elevación en todo el dominio espacial. La Figura 2-8 muestra los resultados de la ecuación (2-1) en cada elemento (pixel) del dominio.

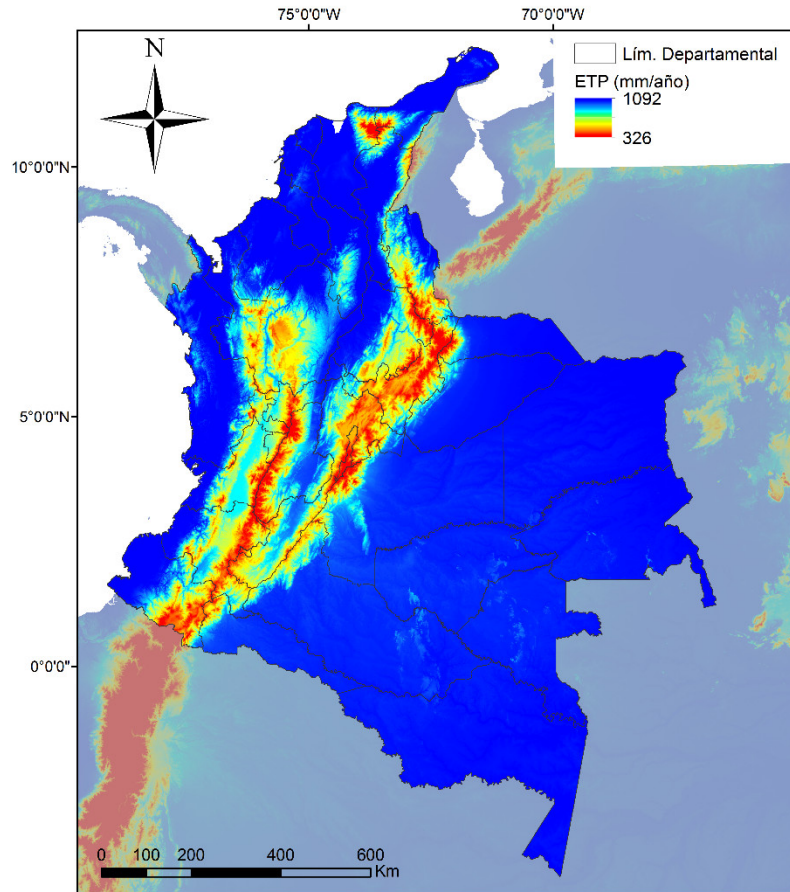


Figura 2-8. Evapotranspiración potencial anual de Colombia.

2.1.7 Temperatura media en superficie

La temperatura media del aire en la superficie terrestre se estima a partir de la elevación según las ecuaciones regionales propuestas por (Chaves C. & Jaramillo R., 1998), (ecuaciones desde la (2-2) hasta la (2-5)) las cuales estiman dicha temperatura en grados centígrados a partir de la elevación en metros sobre el nivel del mar (H) y la región del país dónde ésta se ubique. En la Figura 2-9 se muestran los resultados de la estimación de temperatura media en Colombia.

$$\text{Región Andina: } T = 29.42 - 0.0061 H \quad (2-2)$$

$$\text{Región Caribe: } T = 27.72 - 0.0055 H \quad (2-3)$$

$$\text{Regiones Amazonía y Orinoquía: } T = 27.37 - 0.0057H \quad (2-4)$$

$$\text{Región Pacífico: } T = 27.05 - 0.0057H \quad (2-5)$$

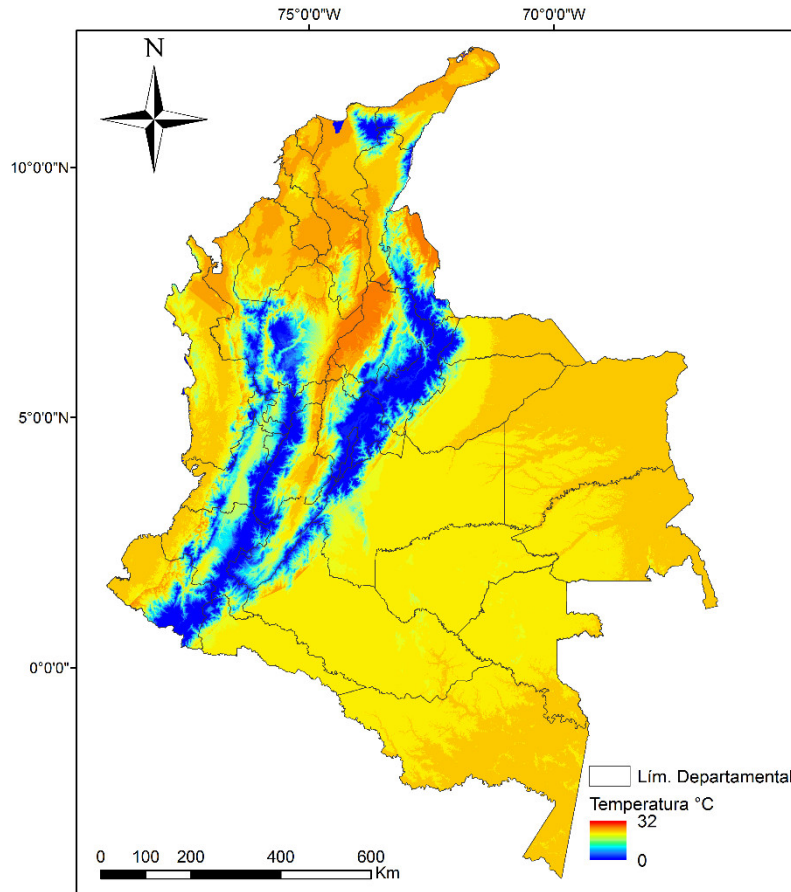


Figura 2-9. Temperatura media del aire en superficie en °C para Colombia.

2.2 Análisis y procesamiento de la información

2.2.1 Tratamiento de las series de caudales

Con el fin de que las curvas de duración construidas con las series de caudal relacionadas en el numeral 2.1.1 abarquen un período hidrológico que se considera lo suficientemente representativo, se depuran todas las estaciones que tengan menos de 15 años en total de operación. De las 849 estaciones de medición revisadas, 723 cumplen con este requisito. También son apartadas del proceso, estaciones que toman registros en brazos fluviales

(ejemplo en Figura 2-11), canales artificiales (ejemplo en Figura 2-12), distritos de riego, ciénagas etc.



Figura 2-10. Estación ubicada en brazo.

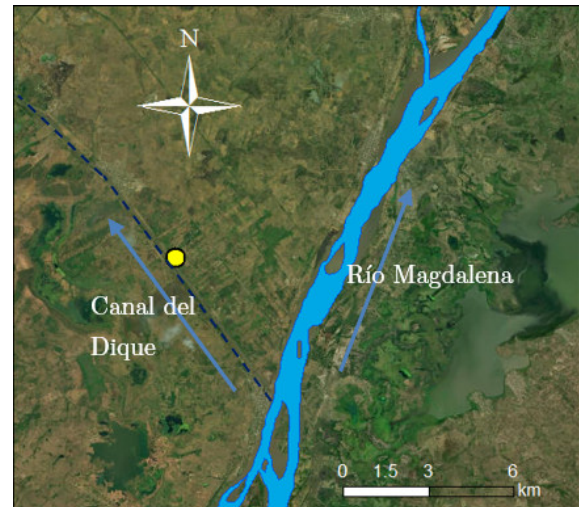


Figura 2-11. Estación ubicada en canal artificial.

Las estaciones de medición ubicadas en las islas de San Andrés, Providencia y Gorgona tampoco fueron tomadas en cuenta para los cálculos pues la regionalización se enfoca en el territorio continental colombiano. Se depuran también estaciones correspondientes a ríos cuyas cuencas no se encuentren en su totalidad dentro del territorio nacional, pues aunque se tienen registros de caudales suficientes, no se cuenta con la información cartográfica para estimar sus características tales como información de clima, geología, cobertura, entre otros. Entre estos, se encuentran ríos como el Orinoco, el Amazonas, el Putumayo entre otros. Luego de realizar todas estas depuraciones, se cuenta con un total de 655 estaciones con series de caudales disponibles para realizar los cálculos, éstas se ubican espacialmente según se observa en la Figura 2-12.

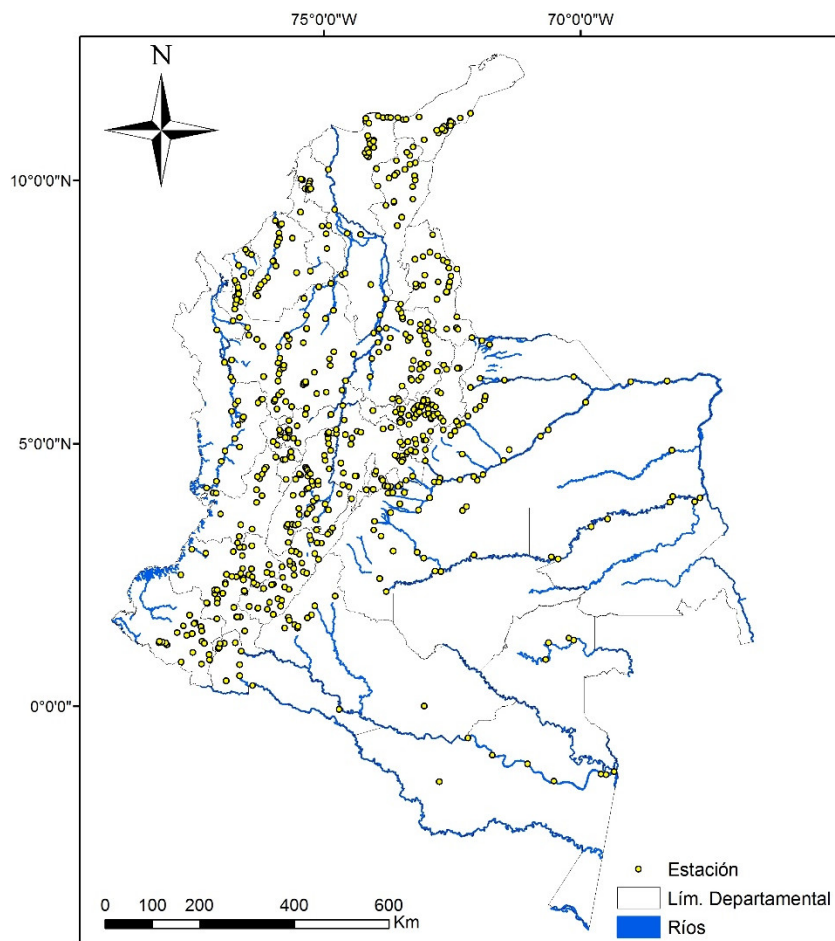


Figura 2-12. Estaciones de medición de caudal definitivas a utilizar, 655 en total.

2.2.2 Procesamiento y manejo del modelo digital de elevación

Al modelo digital de elevación presentado en el numeral 2.1.2 se le realiza una serie de procesos con el fin de adaptarlo para el trazado de las cuencas y la obtención de los diferentes parámetros y atributos topográficos. Los procesos se llevan a cabo en el software de manejo de información geográfica *ArcMap*. Para poder llevar a cabo el trazado y la delimitación de las cuencas, es necesario primero, eliminar todos los sumideros que se encuentren en el dominio de trabajo, para lo cual se utiliza la herramienta *Fill*, posteriormente se utiliza la herramienta *Flow Direction* para generar un esquema bidimensional de direcciones predominantes de flujo.

Se crea el mapa general con la red de drenaje de todas las cuencas mediante la generación preliminar de un mapa de acumulación de flujo (herramienta *Flow Accumulation* de *ArcMap*), éste se reclasifica de tal manera que solo muestre valor en las celdas donde se

acumulen más 740 píxeles, lo que corresponde a un área de drenaje de 6 km². El Instituto Geográfico Agustín Codazzi *IGAC* proporciona también un levantamiento digital en el cual se puede visualizar la red de drenaje de toda Colombia, ésta puede servir como insumo para realizar una comparación y evaluar la precisión de la red de drenaje obtenida a partir del procesamiento del *DEM*.

El mapa de acumulación de flujo, generalmente simula con buenos resultados la red de drenaje en zonas montañosas o de altas pendientes, esto puede observarse en la Figura 2-13 en la cual se compara la red de drenaje suministrada por el *IGAC* y la red producto de procesamiento del *DEM* en una zona montañosa, caso contrario ocurre en zonas planas donde el modelo de elevación es más sensible al ruido y el mapa de dirección de flujo no aproxima con tan buenos resultados la red de drenaje observada (*IGAC*) un ejemplo de esto se muestra en la Figura 2-14.

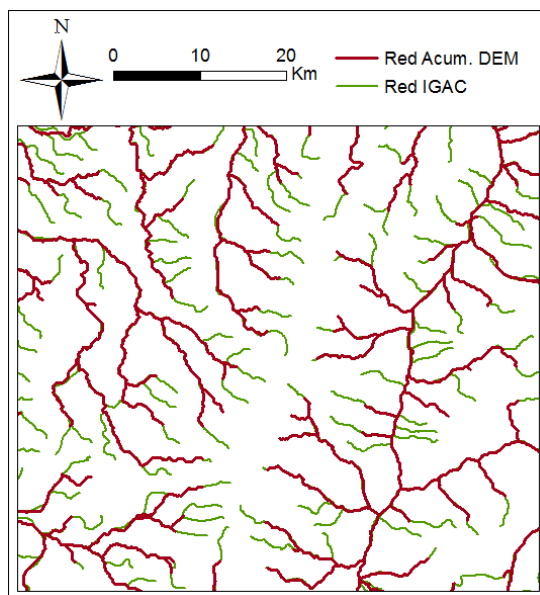


Figura 2-13. Ejemplo de contraste entre redes en zona montañosa.

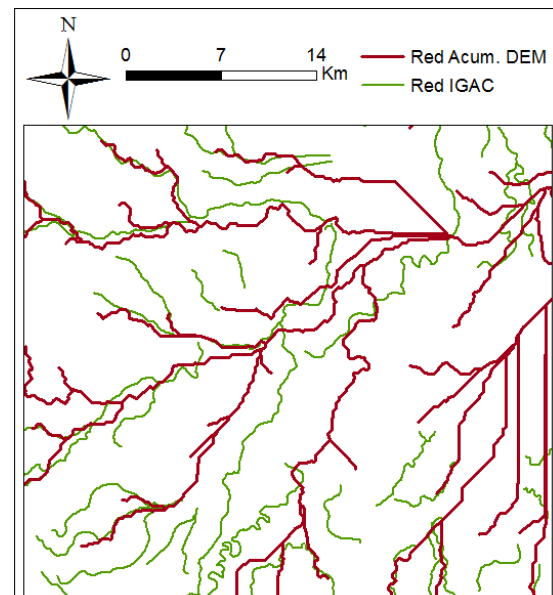


Figura 2-14. Ejemplo de contraste entre redes en zona plana.

Debido a inconsistencias entre la red de drenaje simulada y la red de drenaje observada, los puntos donde se encuentran las estaciones (Figura 2-12) se reubican espacialmente de tal forma que queden ubicados dentro de una celda o píxel perteneciente al drenaje (dentro de la red simulada con el *DEM*), esto permite el trazado de las cuencas en serie utilizando la herramienta *Watershed* en un modelo iterativo. De cada una de las cuencas trazadas y su correspondiente porción espacial del *DEM*, se extraen los atributos: área de drenaje, perímetro, elevación máxima, elevación media, elevación mínima, desnivel, pendiente media

de la cuenca, longitud del cauce principal, elevación máxima, elevación media, elevación mínima y densidad de drenaje.

Se calcula también el índice de compacidad de cada cuenca como parámetro morfométrico, éste compara la forma superficial de la cuenca con la de un círculo mediante una relación entre el perímetro y la raíz del área de drenaje según indica la ecuación (2-6).

$$k_c = 0.28 \frac{P}{\sqrt{A}} \quad (2-6)$$

En la literatura se indica que también algunos percentiles de la curva hipsométrica están relacionados con los caudales característicos de la CDC, por ejemplo (Mohamoud, 2008) utilizó los percentiles 10 y 50 de la curva hipsométrica para estimar caudales normalizados en ríos estadounidenses, mientras que (Sauquet & Catalogne, 2011) utilizaron los percentiles 25, 50 y 75 de la misma como atributos para estimar CDC en cuencas sin información en Francia. A partir de la reclasificación mediante histogramas de las elecciones proporcionadas por el *DEM* y los polígonos correspondientes a las cuencas trazadas, se estiman cuatro percentiles (H_{10} , H_{25} , H_{50} y H_{75}) de las curvas hipsométricas de todas las cuencas.

2.2.3 Cuencas calibradas y cuencas sin calibrar

Se tiene un total 655 cuencas hidrográficas trazadas a partir de las estaciones de medición de caudales diarios del *IDEAM* correspondientes a las estaciones seleccionadas en el numeral 2.2.1. La Figura 2-15 muestra el total de estas cuencas calibradas trazadas dentro del territorio colombiano. Se debe tener en cuenta que los valores de áreas de drenaje se encuentran entre los 6 y los 256.700 km² y que muchas de las unidades de menor tamaño, se ubican espacialmente dentro de las de mayor tamaño, es decir son sub cuencas de éstas. Debido a lo anterior, se ocasiona una superposición o traslape espacial entre los polígonos, lo que dificulta su visualización.

Con el fin de cubrir la totalidad del territorio nacional en cuencas hidrográficas sin que se produzcan superposiciones, es necesario dividir espacialmente el país en unidades hidrológicas, aunque éstas no estén necesariamente instrumentadas, a estos trazados se les denomina en adelante como cuencas sin calibrar. El *IDEAM* dispone también de un mapa de cuencas sin calibrar donde se abarca todo el territorio colombiano denominado Zonificación Hidrológica de Colombia (Salazar-Holguín, 2013). En la Figura 2-16 se

presenta el mapa de cuencas sin calibrar, en total se tienen 394 unidades cuyas áreas de drenaje varían entre los 200 y los 25.000 km².

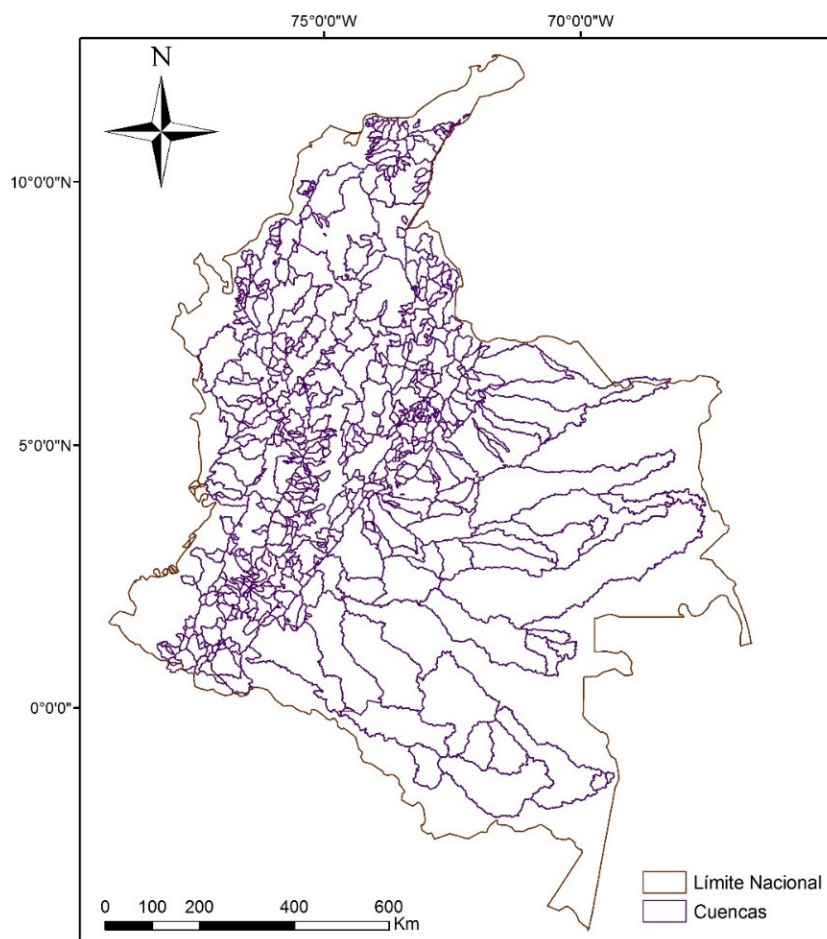


Figura 2-15. Mapa de cuencas calibradas trazadas.

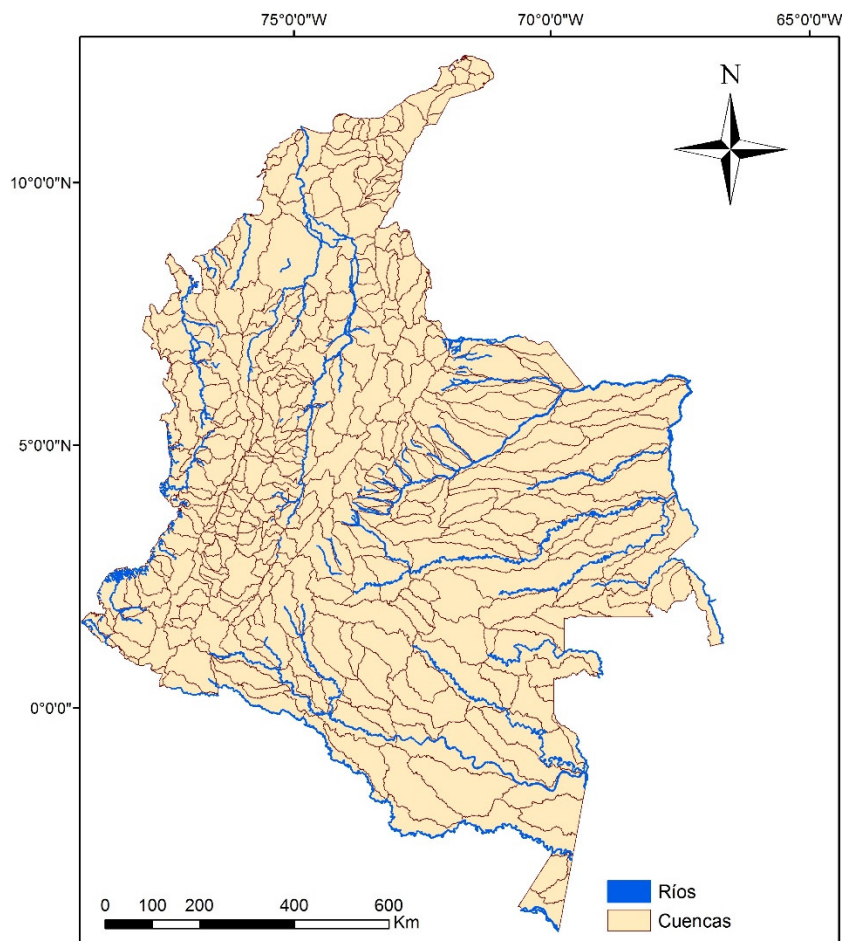


Figura 2-16. Mapa de cuencas sin calibrar en el territorio colombiano. Fuente: (Salazar-Holguín, 2013).

A cada una de las cuencas hidrográficas de Colombia, tanto calibradas como sin calibrar, se le asigna un total de 25 atributos seleccionados entre características referentes a su tamaño, forma, condiciones climáticas del lugar donde se ubica, tipo de cobertura del suelo, geología o material predominante del suelo, y condiciones topográficas. En la Tabla 2-1 se presenta el listado con los 25 atributos o características calculados para cada una.

Tabla 2-1. Resumen de atributos.

Variable	Unidades	Abreviatura
Área de drenaje de la cuenca	km ²	ADre
Perímetro de la cuenca	km	Perim
Índice de compacidad	m/m	Komp
Porcentaje de suelo agrícola	%	%Agr

Variable	Unidades	Abreviatura
Porcentaje de bosque	%	%Bosq
Porcentaje de suelo urbano	%	%Urb
Densidad de falla	(km/km ²)	DenF
Densidad de drenaje	(km/km ²)	DenD
Porcentaje de limos	%	%Lim
Porcentaje de arenas	%	%Aren
Porcentaje de arcillas	%	%Arc
Evapotranspiración potencial media anual	mm/año	ETPa
Elevación máxima	m	E _{max}
Elevación media	m	E _{med}
Elevación mínima	m	E _{min}
Desnivel de la cuenca	m	Desn
Longitud del cauce principal	km	LonC
Pendiente media de la cuenca	%	Pend
Precipitación máxima mensual	mm/mes	P _{máx}
Precipitación media mensual	mm/mes	P _{med}
Temperatura media superficial	°C	T _{med}
Percentil 10 curva hipsométrica	%	H ₁₀
Percentil 25 curva hipsométrica	%	H ₂₅
Percentil 50 curva hipsométrica	%	H ₅₀
Percentil 75 curva hipsométrica	%	H ₇₅

2.3 Análisis de calidad de la información

Mediante la construcción de las curvas de duración con datos observados, es posible identificar valores anómalos, estos son aquellos que difieren en varios órdenes de magnitud, estos datos pueden ser producto de errores en la transcripción o del mismo procesamiento de la información. Se revisa también, que las series de datos no presenten valores muy altos de porcentaje de datos faltantes, en este caso, la mayoría de las estaciones (cerca del 80%) presentan un valor inferior al 30% de este ítem.

Por desfases entre redes de drenaje e imprecisiones en la ubicación de las estaciones, cabe la posibilidad de que alguna cuenca haya sido trazada en un río que no le corresponde. Esto crea la necesidad de revisar la existencia de outlayers o valores anómalos en todo el banco de información, para detectar este tipo de valores, se grafica la dispersión de caudal medio a largo plazo de la serie contra el área de drenaje de las 655 cuencas calibradas como se

muestra en la Figura 2-17. En la figura se observa que en general la relación es muy estrecha y aunque existe dispersión, ésta puede deberse a la alta variabilidad del rendimiento líquido en Colombia, que es la relación entre el caudal medio y el área de drenaje \bar{Q}/A en l/s/km², cuyos valores en teoría deben oscilar entre los 3 y los 200 l/s/km² (Instituto de Hidrología Meteorología y Estudios Ambientales IDEAM, 2014).

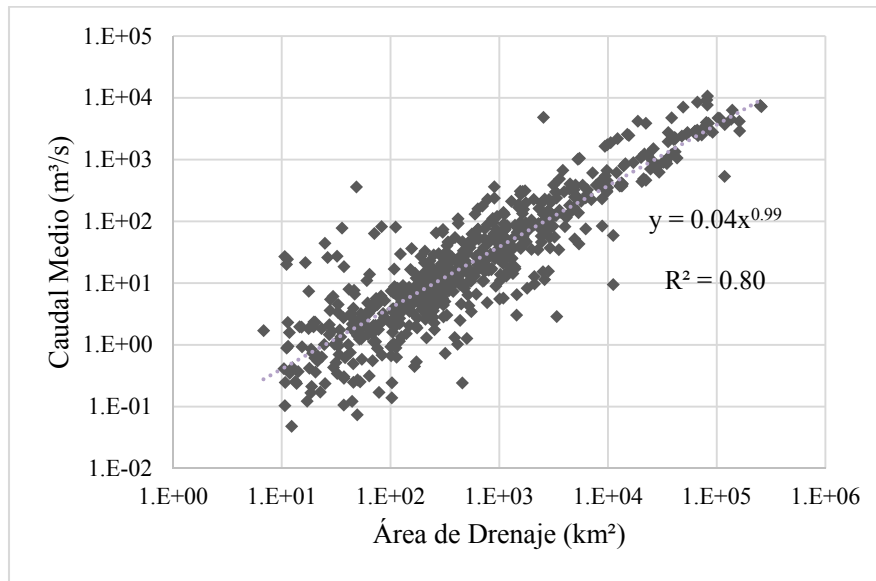


Figura 2-17. Caudal medio a largo plazo vs área de drenaje en las 655 series de caudales.

3 Definición de Regiones Hidrológicas de Colombia

3.1 Introducción

La agrupación de cuencas en regiones hidrológicas tiene gran aplicabilidad con miras a la estimación del régimen de caudales y demás parámetros hidrológicos, ésta se fundamenta en la extrapolación de condiciones desde sitios instrumentados o con información suficiente a otros no calibrados o con información escasa. En este trabajo, se plantea realizar una definición previa de regiones hidrológicamente homogéneas con el fin de realizar la regionalización de la curva de duración de caudales en Colombia. Las regiones hidrológicas se estiman con el propósito de sintetizar la información, de modo que se puedan realizar análisis que apunten a un adecuado manejo del recurso hídrico (García et al., 2017). De antemano cabe aclarar que, los límites de estas regiones hidrológicas, son generalmente, diferentes a las fronteras y divisiones políticas (Comisión Nacional del Agua, 2006), aunque en numerosas ocasiones, tanto el trazado de los ríos como las divisorias de cuencas, delimitan dichos límites administrativos.

La Asociación Mundial para el Agua (GWP, 2008), define la cuenca hidrográfica, como la unidad territorial más apta para la planificación y gestión del recurso hídrico, debido a que en ésta, se integran aspectos físicos, bióticos, políticos y socio económicos alrededor del manejo del agua (Chávez, Jouravlev, & Dourojeanni, 2002). El estudio Nacional del agua de Colombia, para 2014, presenta 5 sub zonas, 41 zonas hidrográficas (Instituto de Hidrología Meteorología y Estudios Ambientales IDEAM, 2014), cada una con diferentes atributos y características fisiológicas y climáticas, de allí surge la necesidad de definir metodologías que permitan sintetizar el estudio de la hidrología en el país.

En Colombia se recomienda utilizar el método de regionalización de parámetros debido a la alta variabilidad climática, topográfica y morfológica que existe en el país, y mediante técnicas estadísticas, estimar variables hidrológicas en zonas con información escasa (Mesa

Sánchez et al., 2003). Es por ello, que se propone en este trabajo una agrupación con enfoque deductivo, el cual consiste en la delimitación de las regiones hidrológicas a partir de atributos relacionados con la respuesta de la cuenca. Por otra parte, el enfoque inductivo propone la delimitación utilizando atributos de la serie temporal (clasificación de caudales) (Olden, Kennard, & Pusey, 2012).

Para el presente trabajo se cuenta con cuencas con información de caudales cuyos tamaños varían desde los 6 hasta los 256.700 km² y una serie de cuencas a agrupar, sin información, que abarca elementos con áreas de aproximadamente entre los 250 y 25.000 km². El objetivo de este capítulo es definir características de las cuencas hidrográficas que sirvan de insumos para la agrupación, por lo cual se busca cuantificar relaciones entre cada uno de los atributos mostrados en la Tabla 2-1 y un set de caudales característicos Q_p de las respectivas curvas de duración. Esto con el fin de realizar un clúster que abarque el territorio colombiano (agrupación de las cuencas sin calibrar).

3.2 Metodología

Se pretende agrupar las cuencas que cuentan con registros de caudales o cuencas calibradas con el fin de establecer parámetros que permitan la agrupación de todas las cuencas sin calibrar según similitud hidrológica, con lo que se obtiene por extensión, la delimitación de regiones hidrológicas del país. Se busca cuantificar relaciones predominantes entre los atributos de las cuencas en general y cada uno de los percentiles característicos Q_p para definir cuál o cuáles variables tienen más peso para la realización de la agrupación.

Se construyen CDC observadas para cada una de las estaciones de medición con registros de caudales diarios preseleccionadas según el numeral 2.1.1. Con el fin de realizar un análisis detallado, abarcando suficientes puntos en cada CDC, se reconstruyen con 30 percentiles Q_p diferentes, los porcentajes de excedencia se eligen de tal forma que se permita la visualización detallada de la curva de duración en los eventos máximos extremos (porcentajes de excedencia cercanos a cero). Los percentiles seleccionados son: Q_{100} , Q_{95} , Q_{90} , Q_{85} , Q_{80} , Q_{75} , Q_{70} , Q_{65} , Q_{60} , Q_{55} , Q_{50} , Q_{45} , Q_{40} , Q_{35} , Q_{30} , Q_{25} , Q_{20} , Q_{15} , Q_{10} , Q_5 , Q_4 , Q_3 , Q_2 , Q_1 , $Q_{0.5}$, $Q_{0.3}$, $Q_{0.1}$, $Q_{0.05}$, $Q_{0.03}$, $Q_{0.01}$. Donde Q es el caudal en m³/s y cada subíndice representa el correspondiente porcentaje de excedencia.

Las CDC reconstruidas de esta forma serán denominadas como CDC reales u observadas, a continuación en la Figura 3-1 se muestra como ejemplo la CDC real de la serie de datos de la estación 11037010 correspondiente al río Quito ubicado en el departamento del Chocó

(área de drenaje de 1568 km² y caudal medio igual a 273 m³/s). En la Figura 3-2 se muestra la CDC real u observada de la estación 21047010 correspondiente al río Magdalena a la altura del departamento del Huila (área de drenaje de 5652 km², caudal medio igual a 213 m³/s).

Se puede observar como una cuenca con menos de la tercera parte del tamaño (río Quito), puede llegar a tener mayor caudal medio y un régimen de descargas mucho más estable o regulado (menor relación entre caudales máximos y caudales medios). Además de cambios en el régimen climático, estas cuencas presentan diferentes usos del suelo, por lo que se evidencia una mayor regulación hídrica en la corriente chocoana. Este es un ejemplo de la alta variabilidad hidrológica que se presenta en Colombia, se espera que estos dos ríos por ejemplo, dadas sus diferentes condiciones, se encuentren en agrupaciones diferentes finalizado el capítulo.

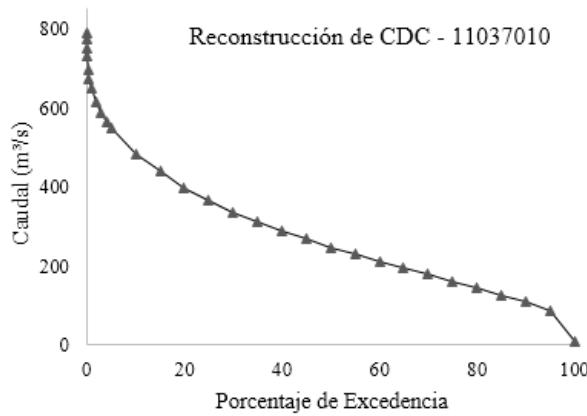


Figura 3-1. CDC real estación 11037010.

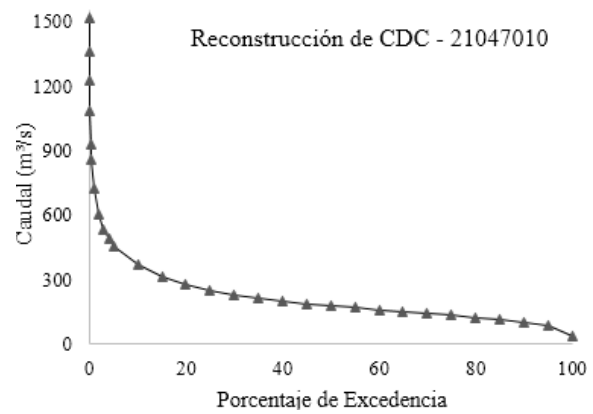


Figura 3-2. CDC real estación 21047010.

Con el propósito de evitar redundancias en los cálculos relacionadas con el tamaño de las cuencas, se procede a realizar tanto una adimensionalización respecto al caudal medio (ecuación (3-1)), como una normalización respecto al área de drenaje de la cuenca (ecuación ((3-2)).

$$Q_p^* [Adim] = \frac{Q_p}{\bar{Q}} \quad (3-1)$$

$$Q_{p \text{ Norm}} \left[\frac{m^3}{s \text{ km}^2} \right] = \frac{Q_p}{A} \quad (3-2)$$

Con base en lo anterior, se presentan los mismos ejemplos de CDC observadas para caudales adimensionales (Figura 3-3 y Figura 3-4) y de caudales normalizados (Figura 3-5 y Figura 3-6).

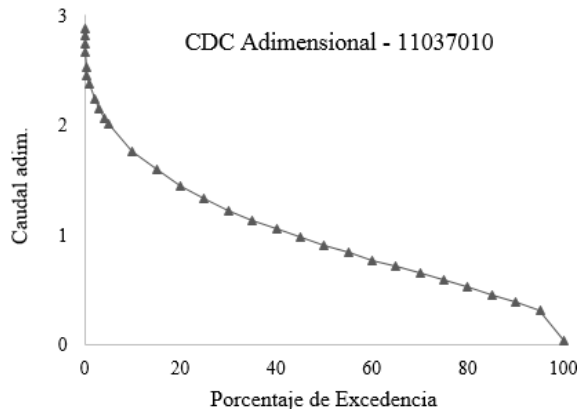


Figura 3-3. CDC adim. Estación 11037010.

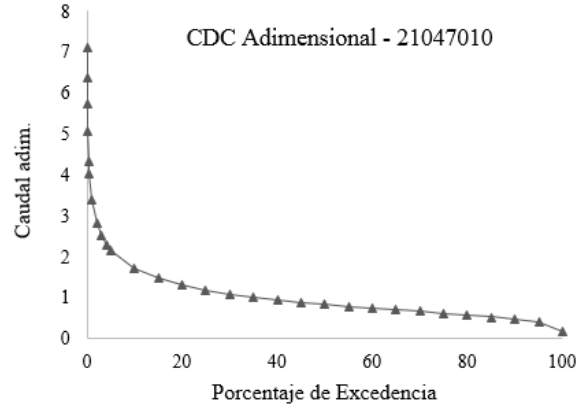


Figura 3-4. CDC adim. Estación 21047010.

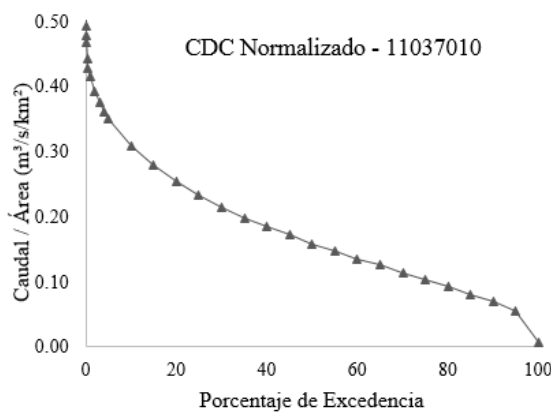


Figura 3-5. CDC Normalizada estación 11037010.

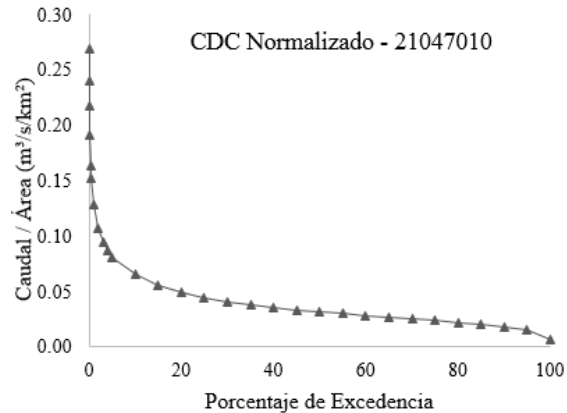


Figura 3-6. CDC Normalizada estación 21047010.

Se observa que la forma de la curva no varía en ambos casos según el proceso de estandarización, el cambio se produce en los valores del eje de las ordenadas. En este trabajo se proyecta una aproximación a la delimitación de diferentes regiones hidrológicas según las diferentes magnitudes de caudales o tramos de la curva de duración (máximos, medios y mínimos), para ello, se realiza un análisis multivariado, con el cual se estudian las relaciones entre un grupo de variables independientes con una o varias variables dependientes a las cuales puedan explicarse características o comportamientos con base en las variables del primer grupo (Anderson, 1958).

Se seleccionan entonces caudales característicos según los diferentes tramos de la curva de duración. (Sauquet & Catalogne, 2011), sugieren que el Q_{10} es un punto de inflexión o quiebre entre los medios y los máximos de la CDC, por lo tanto, percentiles que se encuentren por debajo de éste pueden ser seleccionados para representar el tramo de máximos. Por otro lado, se ha encontrado que el caudal medio (promedio aritmético de toda la serie) se encuentra ubicado generalmente entre los percentiles 30 y 40 (dependiendo del tamaño de la cuenca y del régimen de lluvia), por lo que puede escogerse entre los caudales Q_{40} , Q_{35} o Q_{30} para representar el régimen de medios. Para evaluar si existen relaciones entre los caudales base o de recesión, se evalúan correlaciones R^2 entre los atributos y los componentes con alguno de los percentiles 85, 90 y 95 de la CDC (Q_{85} , Q_{90} y Q_{95}).

La agrupación de cuencas, puede realizarse mediante un análisis de clúster, el cual consta de la selección de grupos dadas unas características o parámetros de agrupación dentro de un conjunto de datos determinado (Wilks, 2011). En este caso el banco de datos a agrupar son tanto las cuencas calibradas (655 unidades) como las cuencas sin calibrar (394 unidades), y las características con las cuales se realiza la agrupación son los 25 atributos climáticos y morfométricos definidos para cada una.

3.3 Modelo de agrupación de cuencas

Se pretende encontrar los atributos que en mayor proporción expliquen o que mejor estén correlacionados con cada uno de los percentiles característicos, se utiliza como medida el coeficiente de correlación lineal R (ecuación (3-3)) con el fin de identificar relaciones existentes entre los diferentes atributos y los valores de las diferentes magnitudes de caudal.

$$R(X, Y) = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}} \quad (3-3)$$

Donde X y Y son las matrices a las cuales se les calcula el coeficiente y \bar{x} , \bar{y} los valores promedio de éstas respectivamente. Se construye una matriz de coeficientes de correlación lineal R entre la matriz de atributos y matriz de caudales adimensionales Q_p^* y matriz de caudales normalizados $Q_{p \text{ Norm}}$. Un resumen con los resultados de las matrices de correlaciones se presenta en el Anexo A.

En la Figura 3-7, Figura 3-8 y Figura 3-9 se muestran a manera de ejemplo, los valores de coeficiente de correlación entre cada atributo y los caudales característicos Q_{90} , Q_{35} y Q_{10} .

respectivamente. Tanto en las figuras mostradas como en el Anexo A, puede observarse que estos valores de R son en su mayoría pequeños o poco significantes, sin embargo, éstos se contrastan posteriormente con los obtenidos en cada clúster luego de haber realizado la agrupación de cuencas, donde se espera encontrar un incremento.

Se trabajan las correlaciones a partir de la matriz de caudales adimensionales ya que de forma generalizada, presentan valores de R más altos que los obtenidos con los percentiles de la matriz de caudales normalizados según el área de drenaje ($Q_p \text{ Norm}$).

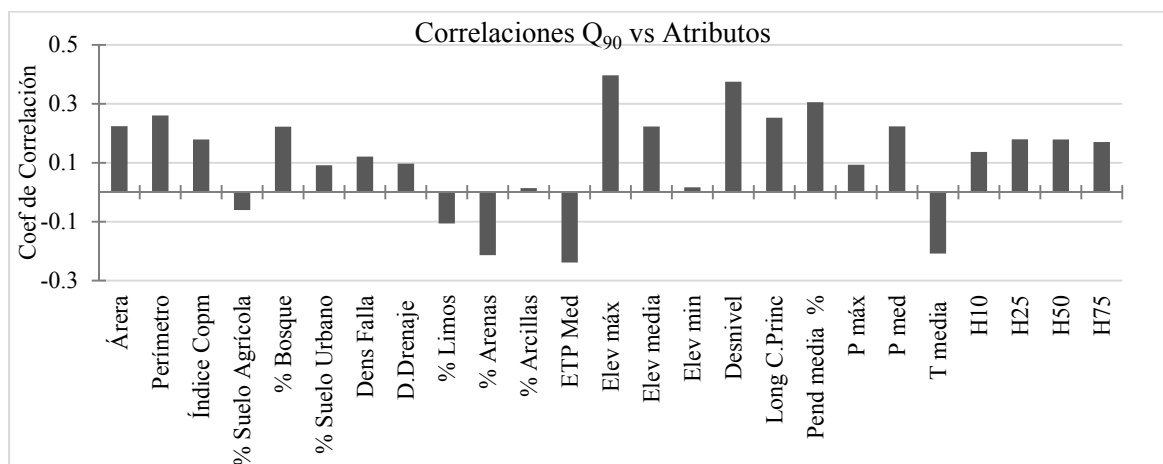


Figura 3-7. Correlaciones Q_{90} adimensional vs Atributos.

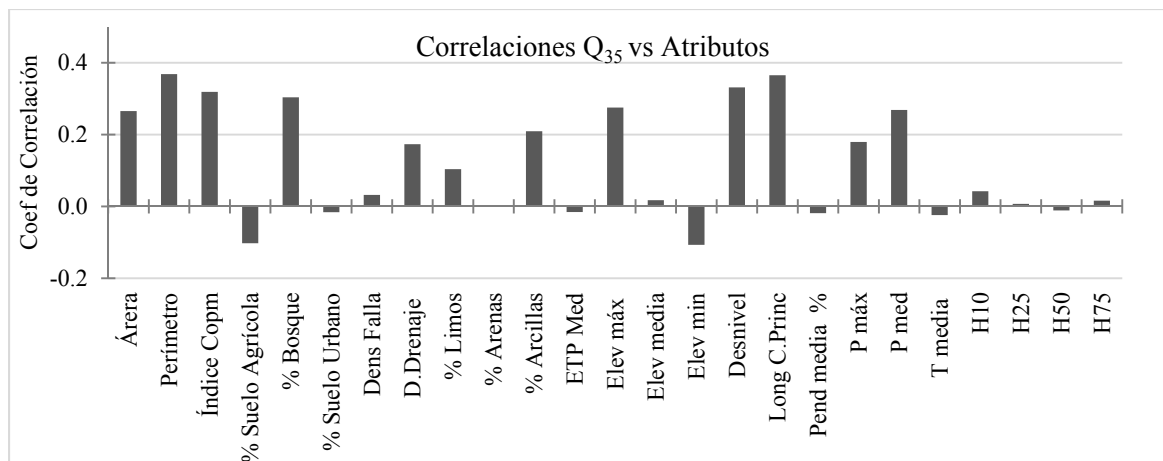


Figura 3-8 Correlaciones Q_{35} adimensional vs Atributos.

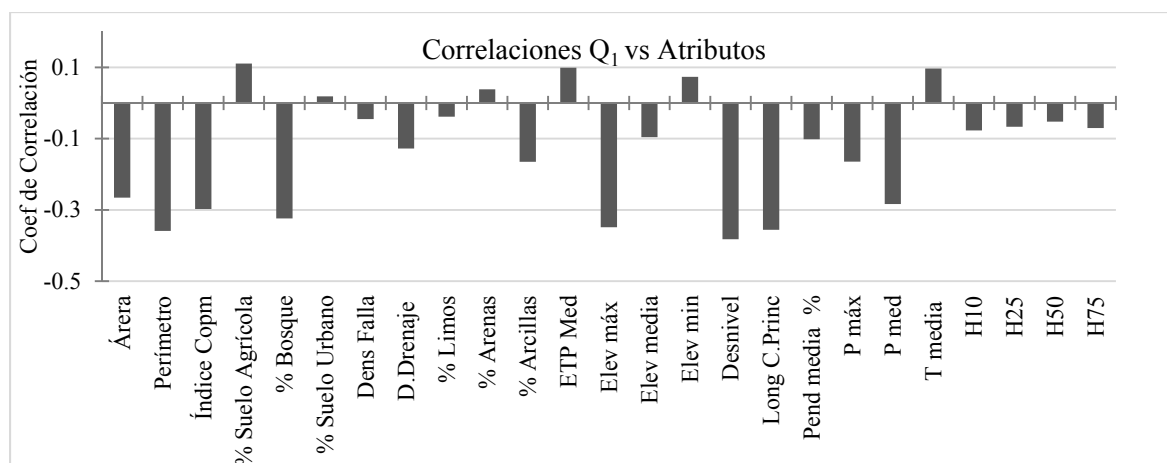


Figura 3-9 Correlaciones Q_1 adimensional vs Atributos.

Existen variaciones de R entre los diferentes percentiles dado un atributo, como se observa en las gráficas anteriores y explícitamente en la Figura 3-10, donde se muestran valores de R variando según porcentaje de excedencia, en este caso se escoge el atributo Q_P vs Desn (m) ya que es uno de los que presenta mayores valores de correlación en el análisis general.

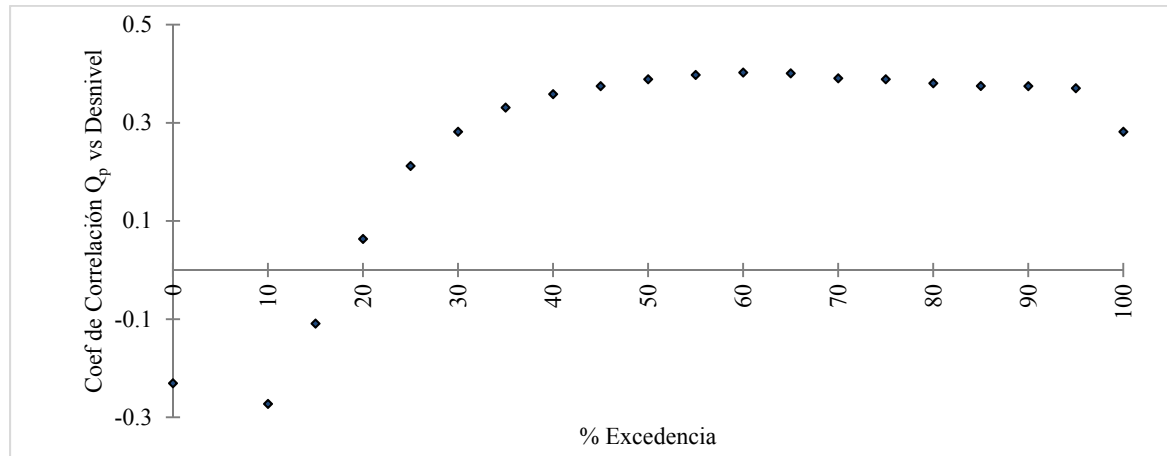


Figura 3-10. Ejemplo de variación percentil a percentil del coeficiente de correlación.

Es necesario listar las variables más influyentes según el anterior análisis en cada uno de los tramos de la CDC. En la Tabla 3-1 se muestran en orden descendente las diez variables de mayor correlación en valor absoluto con los percentiles característicos seleccionados (90, 35 y 1 en porcentaje de excedencia). Se observa que las variables de mayor relevancia en los procesos, según este análisis, no son muy diferentes evaluando en las distintas magnitudes

de caudal. Debido a lo anterior, se opta por descartar el análisis de agrupación y conformación de regiones hidrológicas según magnitud de caudal, esto ocasiona que el análisis de clúster sea más simple, puesto que se realiza una única agrupación.

Tabla 3-1. Atributos seleccionados para cada tramo de la CDC

Máximos Q_1	Medios Q_{35}	Mínimos Q_{90}
Desn	Perím	Emáx
Perím	LonC	Desn
LonC	Desn	Pend
Emáx	Komp	Perím
%Bosq	%Bosq	LonC
Komp	Emáx	ETPmed
Pmed	Pmed	ADre
ADre	ADre	Pmed
%Arc	%Arc	Emed
Pmáx	Pmáx	%Bosq

3.3.1 Selección de número de grupos

Para la conformación de los grupos, se utiliza el algoritmo de *K means clustering* o *k means*, el cual consiste en agrupar mediante dividir M puntos de N dimensiones en K clústers tales que la suma de las distancias euclidianas de los puntos al centroide de su respectivo clúster sea la mínima posible (Hartigan, 1975). La distancia euclidiana viene dada según la ecuación (3-4),

$$d_E(P, Q) = \sqrt{\sum_{i=1}^N (p_i - q_i)^2} \quad (3-4)$$

donde P y Q son dos puntos en un espacio de N dimensiones. Las coordenadas del centroide geométrico en los clúster, corresponden a la media aritmética de cada dimensión. Cada una de las cuencas calibradas es un punto M ($M = 655$), mientras que la dimensión N está dada en este caso por el número de atributos de cada cuenca, ($N = 25$). En la Figura 3-11 se muestra a manera de ejemplo un esquema de clúster en dos dimensiones (“Variable 2” vs “Variable 1”), en el lado izquierdo de la figura se muestran los datos en una única dispersión

sin agrupar, mientras que en el lado derecho de la figura aparecen en diferentes marcadores 3 grupos luego de la aplicación del algoritmo.

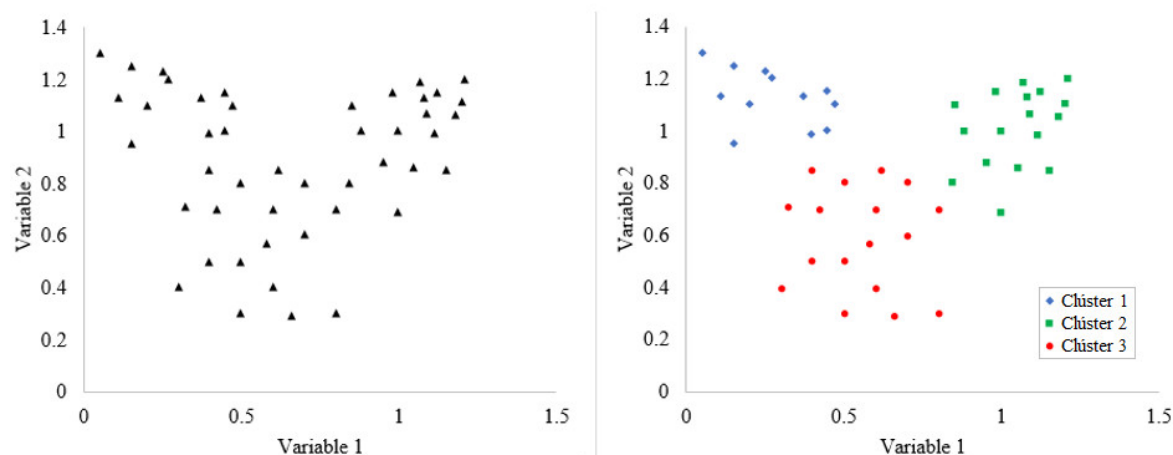


Figura 3-11. Esquema ilustrativo *K means*, izquierda: dispersión sin agrupar, derecha: grupos conformados.

K-means Clustering o *K means* fue utilizado con el fin de agrupar las cuencas según similitud en sus atributos utilizando el software *MATLAB*, el algoritmo requiere como insumos una matriz de datos X (de dimensiones 655 por 25) y un número de grupos o clústers previamente definidos K , se generan K centroides con coordenadas en N dimensiones aleatorias y se calcula la distancia euclidiana de cada punto a su centroide. Se realizan varias iteraciones variando las posiciones de los centroides inicialmente establecidos minimizando las distancias euclidianas medias de cada grupo de puntos a sus respectivos centroides. Como resultado se obtiene un vector S donde se indica el número del grupo al que pertenece cada cuenca, acompañado de un vector D que contiene las distancias de cada punto a su centroide.

Con el fin de determinar cuántos grupos se conforman, se computa el algoritmo *K means* tanto para la matriz de atributos con sus valores estandarizados como para la matriz de atributos con sus valores originales, variando el valor de K (número de grupos) y evaluando la distancia media a los respectivos centroides. Los resultados obtenidos se observan en la Figura 3-12 para los atributos no estandarizados y en la Figura 3-13 para los atributos estandarizados. La estandarización se realiza restando a cada dato el valor de la media aritmética \bar{X} y dividiendo por la desviación estándar σ .

$$X_{est} = \frac{X - \bar{X}}{\sigma} \quad (3-5)$$

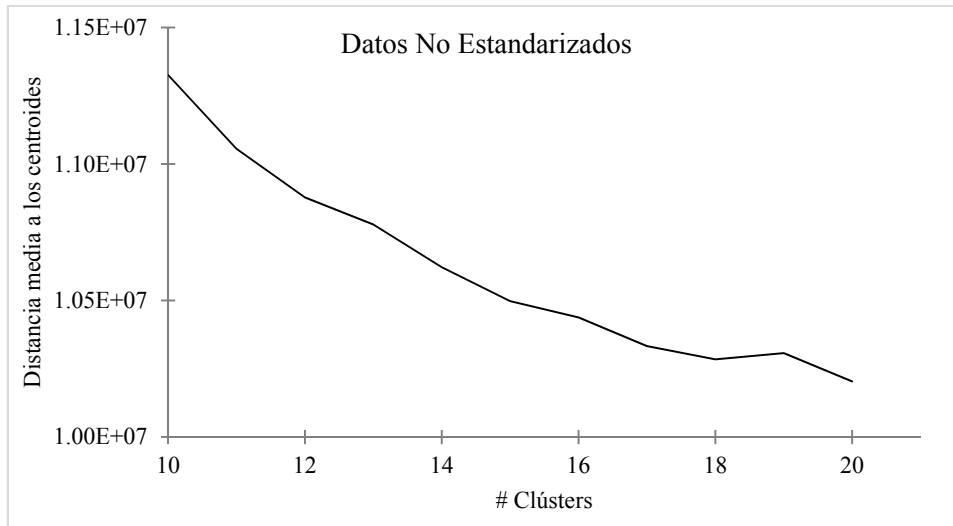


Figura 3-12. Análisis de sensibilidad *K means* con datos no estandarizados.

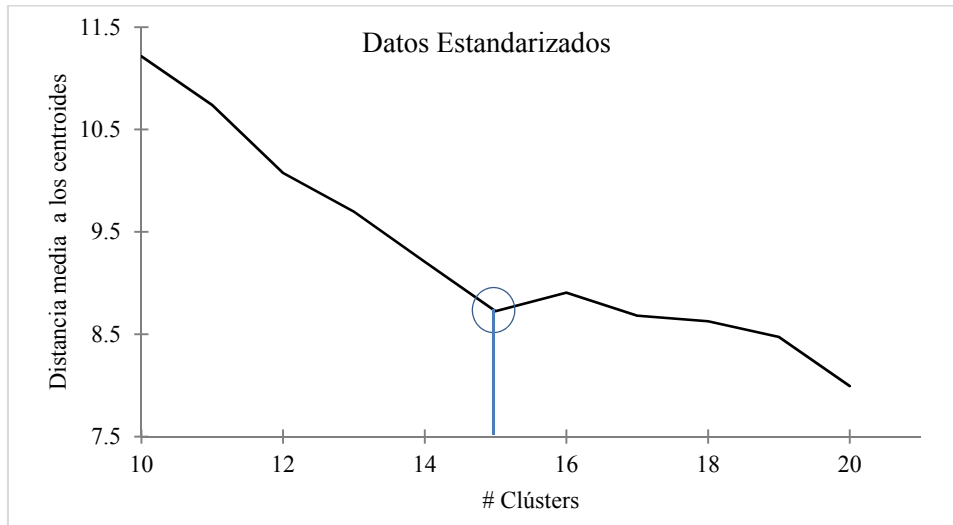


Figura 3-13. Análisis de sensibilidad *K means* con datos estandarizados.

Comparando ambas figuras, se tiene un cambio de pendiente o quiebre más definido en la abscisa 15 de la Figura 3-13 (con los datos estandarizados), cabe resaltar también que las distancias euclidianas presentan gran contraste entre una figura y otra, puede resultar más

sencillo trabajar con las magnitudes de las distancias medias para los datos estandarizados (valores entre 7 y 12) que con las magnitudes de las distancias medias para los datos no estandarizados (del orden de 10^7).

Con base en el anterior análisis de sensibilidad se llega a la conclusión de que es óptimo definir un número de agrupaciones (parámetro K) igual a 15, cabe destacar, que mientras más grupos se conforme, efectivamente la distancia a los centroides tiende a disminuir, pero se hace más compleja la agrupación. (P. L. García et al., 2017), apoyados por juicio experto definieron 16 regiones hidrológicas en el territorio colombiano, por lo que se puede concluir que la literatura reporta un valor cercano, acorde con los resultados de este análisis.

3.4 Aplicación de clúster y resultados

La agrupación de las 655 cuencas calibradas se realiza utilizando el algoritmo de *K means* a partir de los 25 descriptores extraídos para cada una y estandarizados según la ecuación (3-5), y generando los 15 grupos establecidos en el numeral 3.3.1. Se muestran resultados preliminares de la agrupación, número de cuencas por grupo y región hidrológica predefinida por el IDEAM en la Tabla 3-2.

Tabla 3-2. Resumen de agrupación de cuencas calibradas

Clúster #	N° de cuencas	Región Predominante
1	53	Magdalena-Cauca
2	21	Magdalena-Cauca
3	12	Magdalena-Cauca y Orinoquía
4	82	Magdalena-Cauca y Caribe
5	38	Caribe y Orinoquía
6	29	Magdalena-Cauca y Orinoquía
7	16	Magdalena-Cauca
8	62	Magdalena-Cauca
9	26	Orinoquía y Amazonía
10	57	Magdalena-Cauca
11	38	Caribe y Pacífico
12	42	Magdalena-Cauca y Amazonía
13	74	Magdalena-Cauca
14	65	Magdalena-Cauca
15	40	Orinoquía
Suma: 655		

Ya que la hipótesis planteada inicialmente en este trabajo indica que al agrupar un banco de unidades según similitud hidrológica, se obtienen mejores aproximaciones de los diferentes percentiles de la CDC. Se procede a evaluar los valores de los coeficientes de correlación R obtenidos para los diferentes clúster conformados, se seleccionan entonces tres clúster de forma aleatoria (3, 7 y 11) y se presentan los valores de coeficiente de correlación lineal R entre cada uno de los 25 atributos con el aproximado caudal medio (Q_{35}) en las figuras Figura 3-14, Figura 3-15 y Figura 3-16. Cabe anotar que para este análisis se utiliza la matriz con los valores de atributos no estandarizados.

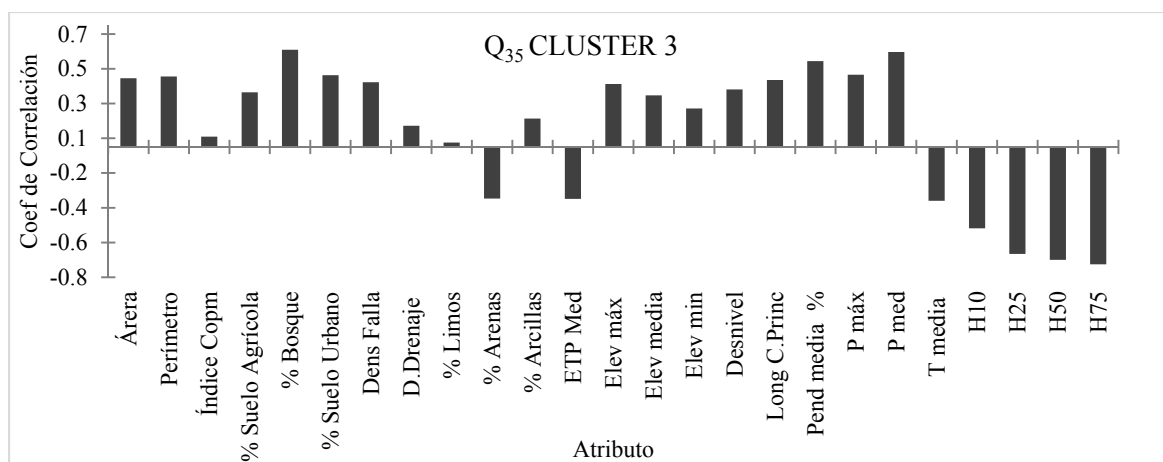


Figura 3-14. Valores de R para Q_{35} vs valor atributo Clúster 3

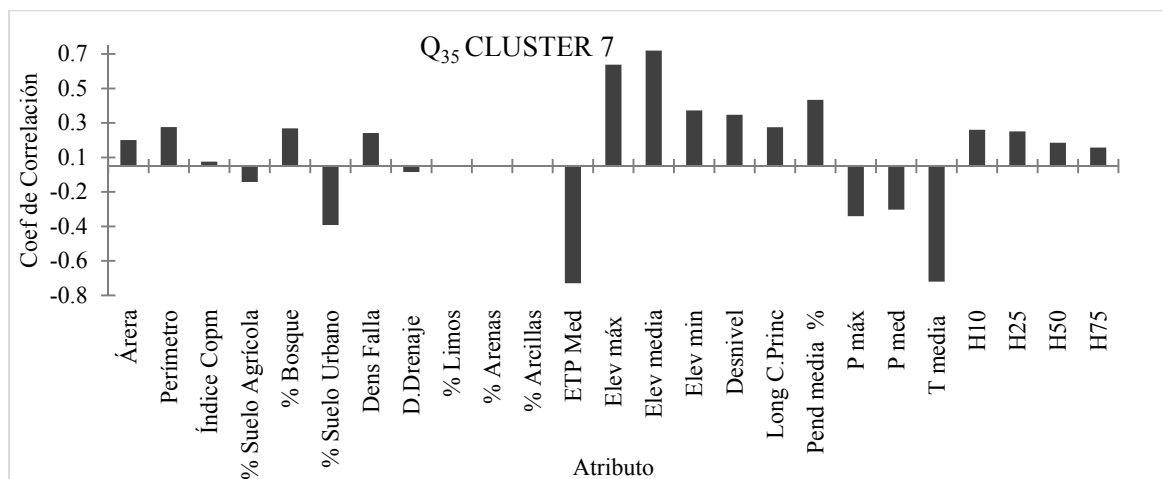


Figura 3-15. Valores de R para Q_{35} vs valor atributo Clúster 7

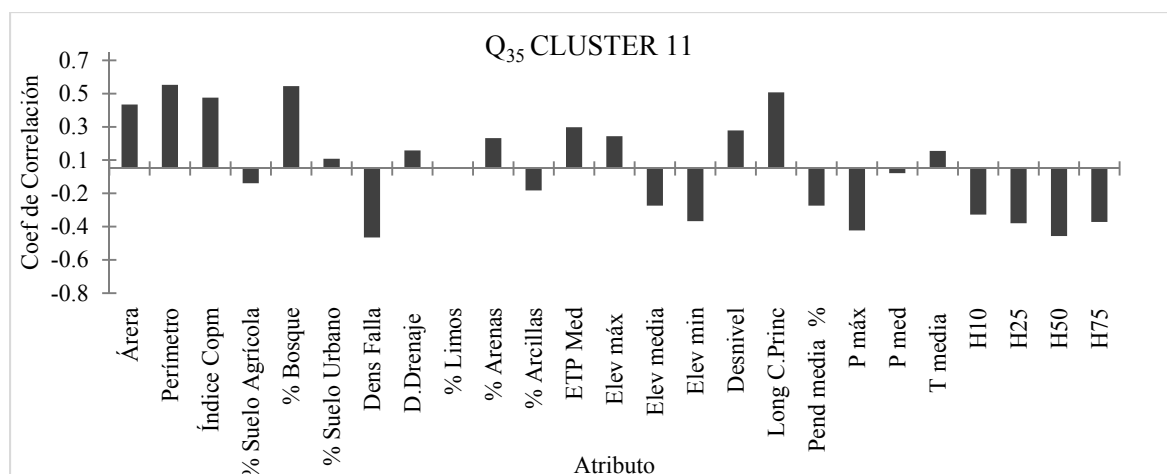


Figura 3-16. Valores de R para Q_{35} vs valor atributo Clúster 11

Las matrices de coeficientes de correlación lineal R (Q_p^* vs atributo en cada clúster) se presentan también en el Anexo A, puede verse tanto allí como en las figuras recién presentadas, que los valores de R son considerablemente mayores en valor absoluto a los presentados por la matriz general de datos (antes de la agrupación). Según esto, la hipótesis: agrupando según similitud en atributos hidrológicos, se obtienen mejores aproximaciones, es verdadera.

3.4.1 Agrupación de cuencas sin calibrar

No puede presentarse un esquema o mapa con las 655 cuencas agrupadas debido a que, como se observa en la Figura 2-15, la mayoría de éstas se encuentran superpuestas entre sí. Uno de los objetivos de este trabajo es mostrar cómo se distribuyen las regiones hidrológicas dentro del territorio colombiano, por lo tanto, para realizar el esquema, se hace necesario realizar la misma agrupación a un grupo de cuencas que abarquen toda el área del país sin que se presenten superposiciones, no es necesario que éstas cuenten con registros de caudales.

Utilizando entonces las cuencas sin información de caudales mostradas en la Figura 2-16, cuyos atributos fueron extraídos de la misma manera y en la misma escala, se procede a realizar la agrupación de las 394 cuencas sin calibrar.

Con el fin de garantizar que las cuencas sin calibrar se ubiquen dentro de los 15 clúster originales (los obtenidos con las cuencas calibradas), se calcula la distancia euclidiana de cada una de éstas a los centroides geométricos, posteriormente se realiza un filtro que

indique cual es la menor distancia de cada cuenca y se asocia el correspondiente centroide con el clúster definitivo. Así se tiene la totalidad del territorio colombiano clasificado en 15 regiones hidrológicas. Esto puede servir como insumo para un futuro estudio de optimización en la extrapolación de parámetros hidrológicos o bien para la estimación del régimen de caudales con la metodología descrita en este trabajo. En el Anexo B, se presentan los centroides geométricos de los clúster con los cuales se realiza la agrupación.

3.5 Análisis de resultados

El mapa de Colombia clasificado según regiones hidrológicas se muestra en la Figura 3-17, mientras que en la Tabla 3-3 se describen los grupos según la región que ocupan espacialmente, se muestra la región o sub región en la cual el grupo predomina. Se hace evidente, que debido a la alta heterogeneidad hidrológica en Colombia se presenten cuencas de un mismo grupo espacialmente distantes.

Tabla 3-3. Resumen de agrupación de cuencas sin calibrar.

Clúster	N° de cuencas	Región Predominante
1	37	Centro de la Orinoquía
2	0	Ríos Cauca y Magdalena (no se ubica en el mapa)
3	4	Pacífico y Bajo Porce
4	38	Guajira, sur de la Amazonía y Orinoquía
5	85	Presente en todas las regiones
6	7	Cordillera central
7	5	Valle del Cauca y Río Medellín
8	21	Cordilleras central y occidental,
9	18	Amazonía y Orinoquía
10	57	Magdalena - Cauca y Caribe
11	24	Pacífico
12	22	Pacífico y Bajo Nechí
13	21	Cordillera occidental y Alto Magdalena
14	8	Valle del Cauca y Boyacá
15	47	Alto y Medio Magdalena y Norte de Antioquia
Suma: 394		

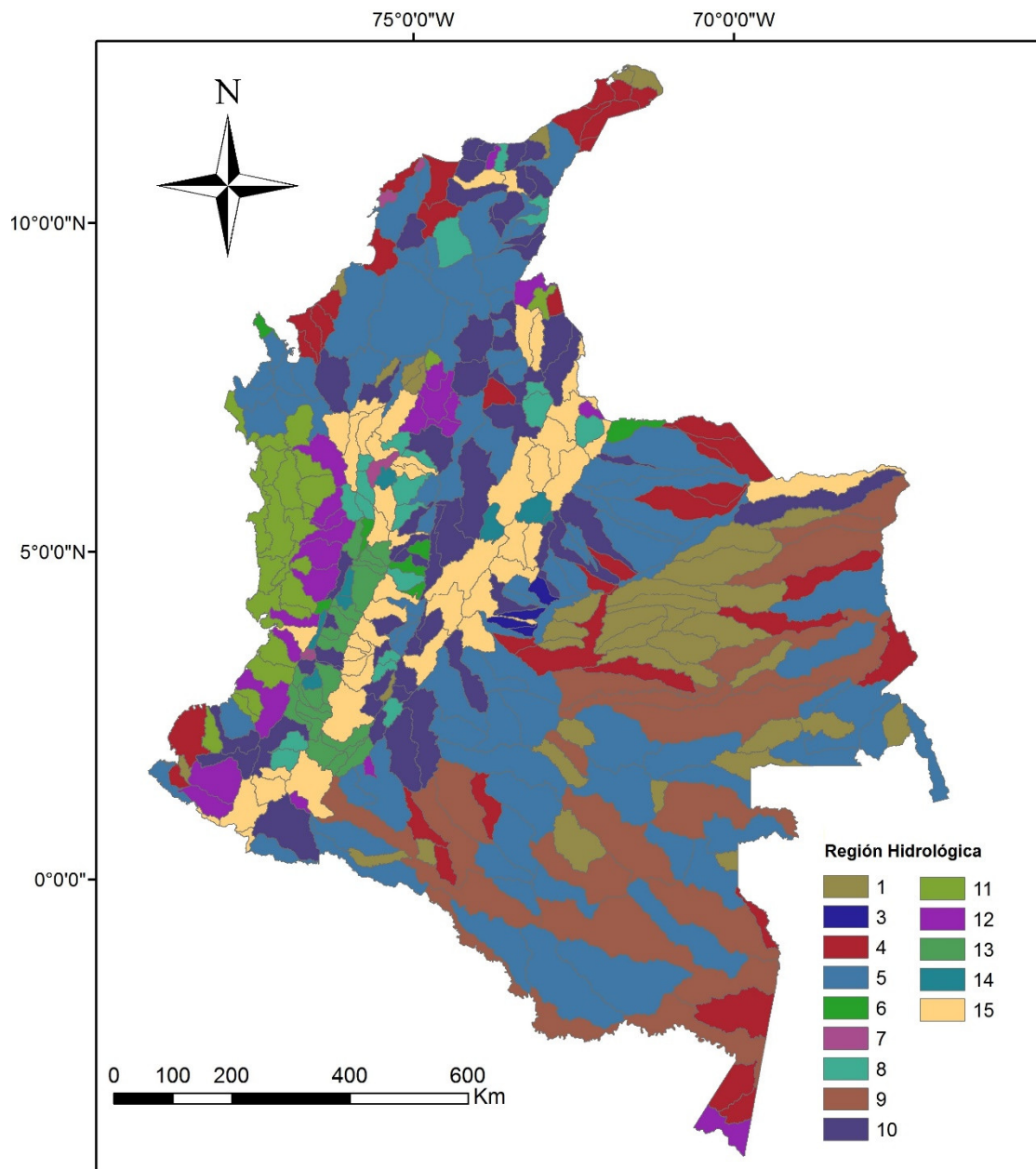


Figura 3-17. Esquema de las Regiones Hidrológicas de Colombia.

Dentro del grupo definido como clúster # 2 no se ubica ninguna de las cuencas sin calibrar, por lo que ésta región hidrológica no aparece presente en el mapa de la Figura 3-17, se seleccionan las cuencas calibradas que conforman el clúster original y se encuentra que todas pertenecen a los ríos Cauca y Magdalena con valores de sus correspondientes áreas de drenaje entre los 30.000 y los 250.000 km², por lo que ninguna de las unidades aquí

agrupadas encaja en estas características. En la Figura 3-17 se hace evidente una alta heterogeneidad hidrológica para Colombia, estos resultados muestran que cuencas ubicadas en sub regiones diferentes, incluso lejanas, pueden llegar a tener comportamientos hidrológicos similares.

A continuación, se presentan diversos mapas referentes a los resultados de agrupación, inicialmente se muestra la distancia euclidiana de cada una de las cuencas al centroide de su respectivo clúster en la Figura 3-18, en ella se observa cuales cuencas del país quedaron con una distancia mayor respecto a los demás elementos de su grupo (cuencas en color azul). Dentro del dominio espacial, se presentan distancias euclidianas considerables que pueden representar cuencas outlayers de un clúster determinado, éstas predominan en el sur de la Amazonía y en la Orinoquía nororiental. Sin embargo, las cuencas mostradas en color verde predominan en el resto del país, por lo que se puede entender como una buena confiabilidad en los resultados.

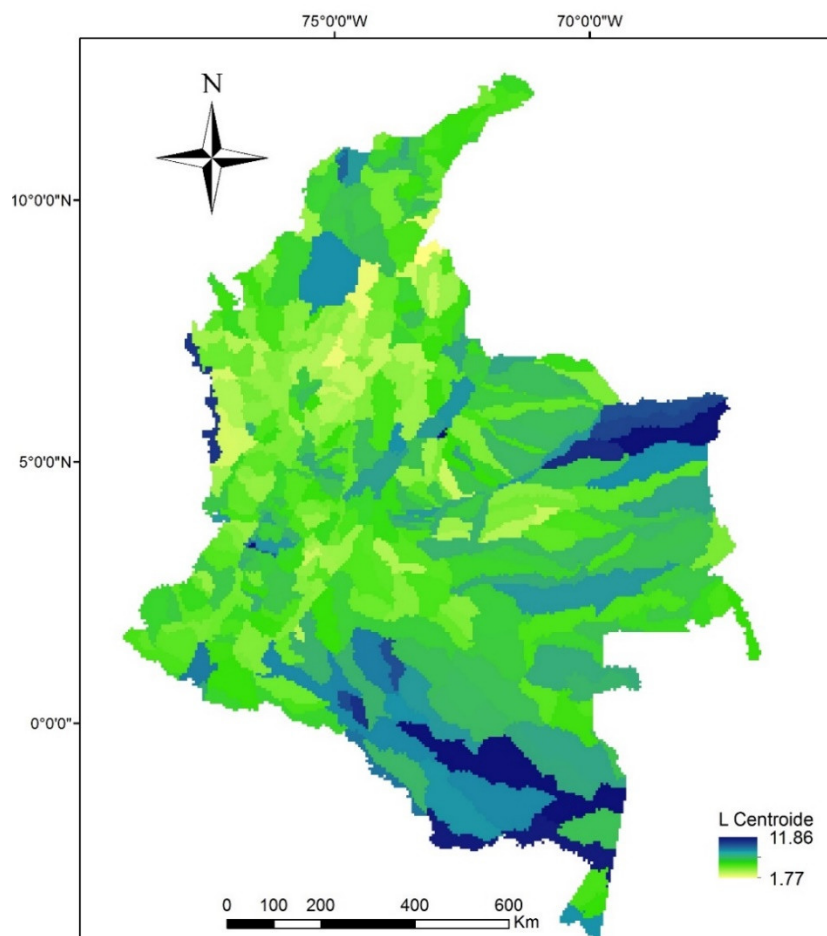


Figura 3-18. Distancia euclidiana al centroide del clúster.

Con el fin de comprobar la alta variabilidad espacial o heterogeneidad de características en las regiones obtenidas, se procede a realizar el mismo análisis espacial de clúster pero aplicado a una sola variable (unidimensional). Se seleccionan variables que hayan representado altos valores de R en el análisis de correlaciones: precipitación media de la cuenca (Figura 3-19), pendiente media (Figura 3-20), evapotranspiración potencial media anual (Figura 3-21) y porcentaje de bosque (Figura 3-22). Cada uno de los clúster individuales se obtiene utilizando el mismo algoritmo *K means* agrupando los datos también en 15 grupos pero teniendo en cuenta una única variable de entrada. Al lado derecho de cada figura, se muestra el valor del atributo seleccionado en todo el dominio, se aclara que no se seleccionan variables que dependen directamente del tamaño de la cuenca (por ejemplo: área de drenaje y perímetro de la cuenca).

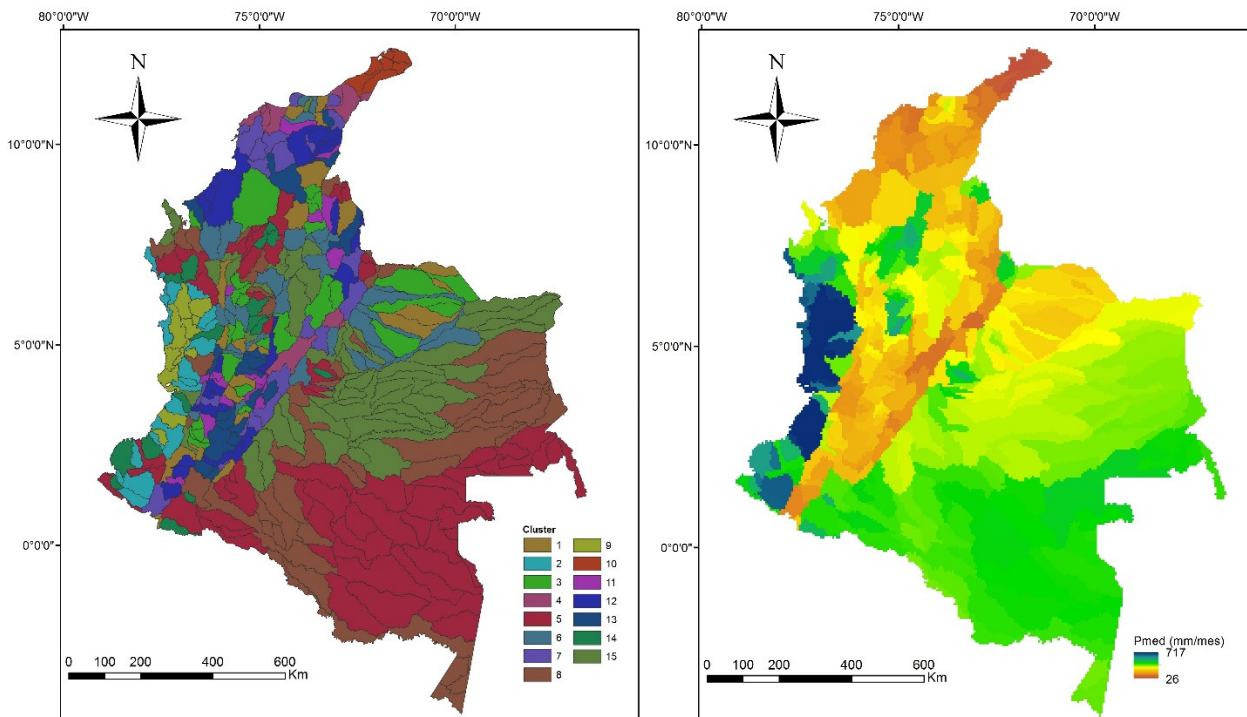


Figura 3-19. Agrupación de cuencas según precipitación (izquierda) y valor de precipitación media de la cuenca (derecha).

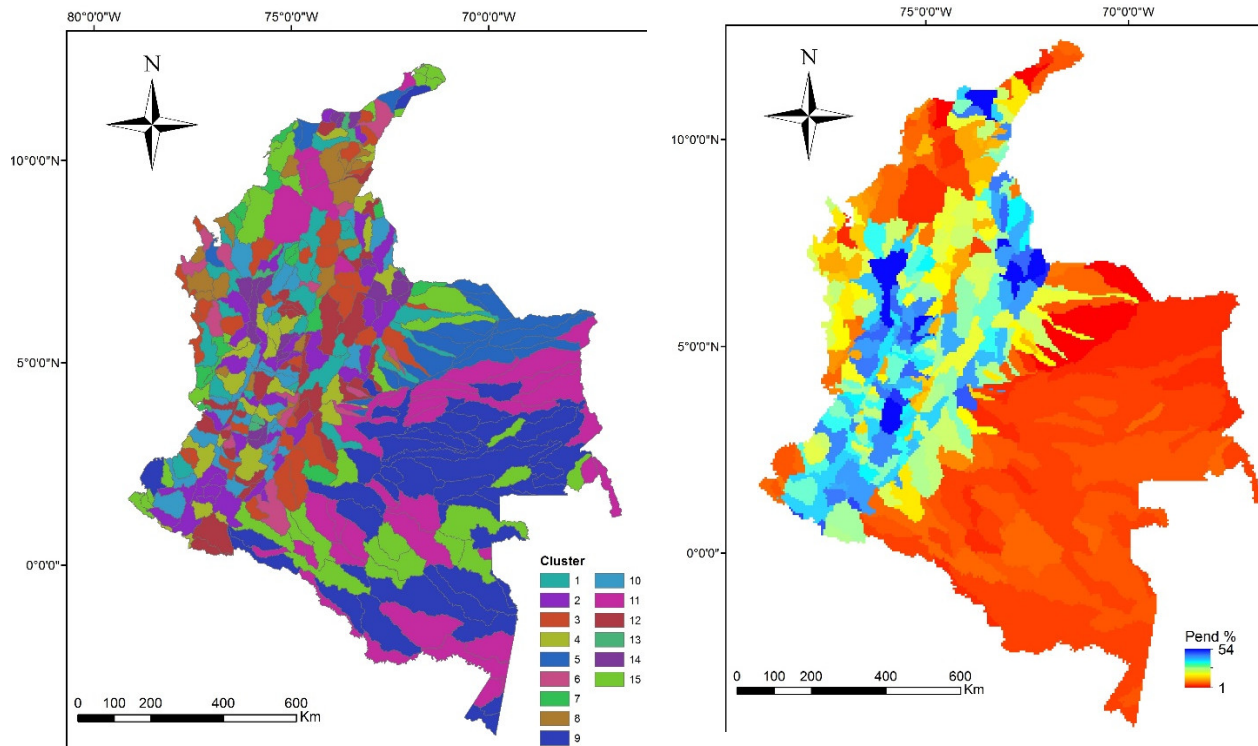


Figura 3-20. Agrupación de cuencas según pendiente (izquierda) y valor de pendiente media de la cuenca (derecha).

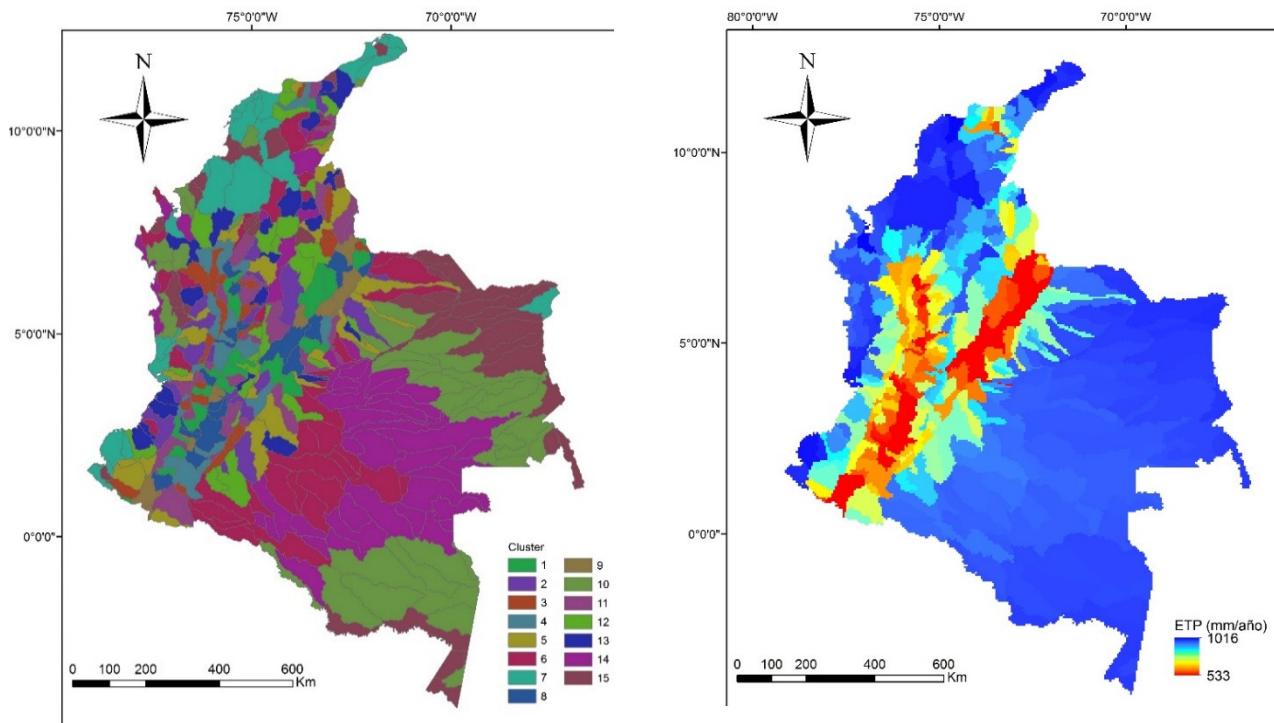


Figura 3-21. Agrupación de cuencas según ETP (izquierda) y valor de ETP media de la cuenca (derecha).

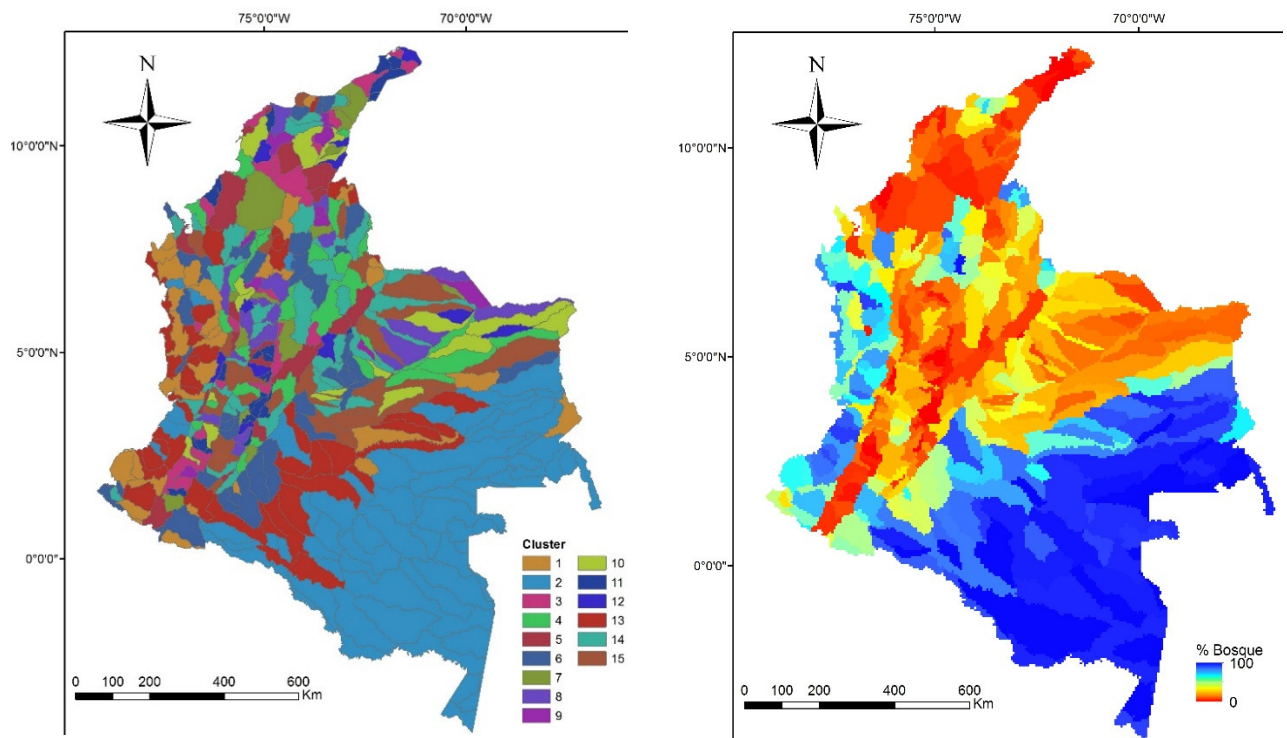


Figura 3-22. Agrupación de cuencas según % de bosque (izquierda) y valor de % bosque de la cuenca (derecha).

Lo anterior puede explicar, la alta heterogeneidad y discontinuidad espacial en las regiones hidrológicas de Colombia, cada una de las variables presentadas y sus respectivos clúster soportan la variabilidad mostrada en el mapa de regiones hidrológicas general de la Figura 3-17.

Según las matrices de correlaciones, generalmente se obtienen valores más altos de R en los caudales medios (Q_{40} , Q_{35} , Q_{30}) vs atributos, esto puede ser una muestra o una evidencia de que los fenómenos que intervienen en el régimen medio de descargas están directamente relacionados, y con mayor claridad, con los 25 atributos preseleccionados o con la mayoría de ellos.

Observando la Figura 3-10, se observa una clara variabilidad percentil a percentil, o entre frecuencias, de los valores de correlación entre diferentes variables, aunque solo se presenta el análisis para una (desnivel) en la realización del trabajo se observó que las demás tienen un comportamiento muy similar. Esto puede contribuir a soportar la hipótesis de que son

distintas las variables y los procesos que intervienen según las magnitudes de descargas, y que por lo tanto valdría la pena realizar agrupaciones hidrológicas pensando en diferentes tramos de la curva de duración. Sin embargo, por efectos de la configuración del clúster ejecutado en este trabajo, no es posible seleccionar diferentes atributos para cada una de las magnitudes de caudal (máximos, medios y mínimos) pues son muy similares las correlaciones entre éstos y los caudales característicos (Q_{90} , Q_{35} y Q_1), por lo que se trabaja en adelante con una única agrupación, sin embargo, queda como interrogante para futuras investigaciones, una redefinición o profundización en este ítem.

4 Regionalización de Curvas de Duración de Caudales

4.1 Introducción

El objetivo principal en este capítulo es estimar la CDC regional para cada uno de los clúster hidrológicos obtenidos en el capítulo anterior (3), se propone un análisis con regresiones lineales múltiples que permitan calcular percentiles característicos tales que, interpolados generan una CDC sintética adimensional estimable a partir de los atributos señalados en la Tabla 2-1.

Según la hipótesis planteada a priori en este trabajo, se lograrán mejores resultados habiendo separado el dominio territorial en regiones hidrológicamente homogéneas, por ello, se propone comparar los valores de R de los diferentes percentiles entre cada uno de los 15 clúster y los valores de R de cada percentil de las CDC estimadas para cada una de las regiones predefinidas por IDEAM (Instituto de Hidrología Meteorología y Estudios Ambientales IDEAM, 2014) Caribe, Magdalena – Cauca, Orinoquía, Amazonía y Pacífico.

Se estiman CDC adimensionales, ya que los distintos caudales característicos de estas presentan mejores coeficientes de correlación y por lo tanto se asume que se obtienen mejores aproximaciones que estimando caudales normalizados según el área de drenaje. Al estimarse el régimen de caudales adimensional, pueden encontrarse los caudales de la curva de duración en m^3/s realizando el producto de cada uno de los Q_p^* por el caudal medio a largo plazo de toda la serie. En caso de no contar con serie hidrológica (puesto que uno de los principios u objetivos del trabajo es estimar el régimen de caudales donde no se tiene información o se cuenta con muy pocos registros) puede estimarse el caudal medio de la corriente mediante la aplicación de balance hídrico a largo plazo:

$$\bar{Q} = C (\bar{P} - \overline{ETR}) Adre \quad (4-1)$$

siendo \bar{Q} el caudal medio de balance a largo plazo, \bar{P} el valor medio histórico de precipitación anual (mm/año), \overline{ETR} el valor medio histórico anual de la evapotranspiración real (mm/año) y $Adre$ el área de drenaje de la cuenca en km^2 , C corresponde a una constante de conversión de unidades de $\text{mm km}^2/\text{año}$ a m^3/s con valor igual a $3,17 \times 10^{-5}$. El modelo de estimación de caudales se realiza según la metodología propuesta por (Mohamoud, 2008) en la cual se generan ecuaciones regionales para 15 percentiles característicos de la CDC. Se selecciona una serie de caudales por cada una de las regiones hidrológicas con la cual se realiza la validación de los resultados arrojados por el modelo, las series seleccionadas no son tenidas en cuenta en el proceso de calibración.

4.2 Modelo de estimación de caudales.

En cada una de las Regiones Hidrológicas de Colombia se estiman curvas de duración a partir de los atributos calculados para cada una de las cuencas en las cuales se tiene registro de caudales, se pretende entonces seleccionar diferentes percentiles característicos Q_p a los cuales se les realiza la estimación, posteriormente, separar las estaciones que a utilizar en la validación de los resultados y que no hacen parte de la calibración del modelo, una por cada clúster y una por sub región definida por IDEAM, aplicar la regresión lineal múltiple (RLM) y finalmente realizar el análisis de los resultados.

4.2.1 Selección de percentiles característicos

Aunque se cuenta con 30 percentiles de las curvas de duración observadas adimensionales, por motivos de simplicidad y facilidad a la hora de realizar la estimación, se trabajan 15 percentiles característicos similares a los propuestos por (Mohamoud, 2008). Se escogen los percentiles para Q_p^* : Q_{100} , Q_{90} , Q_{80} , Q_{70} , Q_{60} , Q_{50} , Q_{40} , Q_{35} , Q_{30} , Q_{20} , Q_{10} , Q_5 , Q_1 , $Q_{0.5}$ y $Q_{0.1}$. Se asume que con estos, puede ser suficientemente diciente la reconstrucción una curva de duración.

4.2.2 Selección de estaciones para validación del modelo

De cada una de las regiones hidrológicas, se selecciona una estación de caudales de manera aleatoria, las cuales no son incluidas dentro del proceso de calibración y que son utilizadas para la posterior validación de los resultados. Se muestra a continuación en la Tabla 4-1 las estaciones asignadas para validar las CDC de cada grupo. Se seleccionan estaciones de diversos lugares del país con el fin de que la validación abarque variadas ubicaciones del

dominio espacial, así mismo, en la Tabla 4-2 se muestran las estaciones seleccionadas para validar los resultados de estimación de CDC en cada sub región hidrológica previamente definida por el IDEAM.

Tabla 4-1. Estaciones de validación para clústers

Clúster	Estación	Región
1	11077020	Caribe
2	23097040	Magdalena
3	35017070	Orinoquía
4	21227010	Magdalena
5	13017010	Caribe
6	35027020	Orinoquía
7	21207960	Magdalena
8	51027020	Pacífico
9	42067010	Amazonía
10	15017010	Caribe
11	32077100	Orinoquía
12	23057010	Magdalena
13	21017020	Magdalena
14	35027150	Orinoquía
15	21197030	Magdalena

Tabla 4-2. Estaciones de validación para sub regiones

Región	Estación
Caribe	13047040
Magdalena	21147080
Orinoquía	35087010
Amazonía	44117010
Pacífico	52027030

4.2.3 Insumos para la calibración del modelo

Para la aplicación de las regresiones, es necesario conformar 15 matrices de atributos (una por cada clúster) con dimensiones $n \times 25$, siendo n el número de estaciones a utilizar en la calibración (habiendo descontado las estaciones destinadas a la validación) y los 25 valores de los correspondientes atributos (con valores sin estandarizar). A este grupo de matrices se les denomina como X . Como variable dependiente se tienen 15 matrices (número de clúster) de $n \times 15$ (número de percentiles característicos a estimar) con los valores de los caudales adimensionalizados respecto al caudal medio (Q_p^*), a esta familia de matrices se les denomina Y , en este caso los datos tampoco están estandarizados.

4.2.4 Regresión lineal múltiple

La regresión lineal múltiple o *RLM* difiere de la regresión lineal simple en que no se analiza la influencia de una sino de varias variables explicativas X en otra variable dependiente Y (Rojo Abuín, 2007), la forma general de una *RLM* está dada por:

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n + \beta \quad (4-2)$$

siendo y cada uno de los percentiles de caudal adimensional Q_p^* , α_i representa el valor del coeficiente de cada término, x_i se refiere a cada uno de los atributos que participan de la regresión, por lo que n este caso toma valores entre 2 y 25, por último, β es el intersección con el eje de las ordenadas o término independiente.

En este caso de aplicación se pretende realizar esta misma regresión, pero en su forma potencial como una multiplicatoria, por lo tanto se aplica una transformación logarítmica a las matrices de datos resultando:

$$\begin{aligned} \log(y) &= \alpha_1 \log(x_1) + \alpha_2 \log(x_2) + \dots + \alpha_n \log(x_n) + \log(\beta) \\ \log(y) &= \log(\beta x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}) \\ 10^{\log(y)} &= 10^{\log(\beta x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n})} \\ y &= \beta x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n} \end{aligned} \quad (4-3)$$

La ecuación (4-3) representa la forma final de la regresión múltiple, en el caso de la regresión potencial se tiene una multiplicatoria de n términos x_i elevado cada uno por su respectivo exponente α_i y acompañados por un coeficiente β . Comparando esta expresión con la ecuación (4-2), el coeficiente β resulta del logaritmo del término independiente, los exponentes α_i son los mismos coeficientes y los valores tanto de x_i como de y son el logaritmo de las matrices originales.

Para realizar este cálculo, tanto los valores de Q_p^* como de x_i deben ser diferentes de cero pues anularía el valor resultante por tratarse de una multiplicatoria, en caso de que alguno lo sea (pues es posible que algunos atributos tomen el valor de cero) se reemplaza por algo cercano como 1×10^{-4} .

Se apunta a realizar primero el cálculo de las regresiones con dos variables independientes cada una, se cargan entonces las 40 matrices de datos al *MATLAB* correspondientes a 15 de atributos y 15 de caudales adimensionales (una por cada clúster) además de 5 de

atributos y 5 de caudales adimensionales (una por cada sub región) y dentro del algoritmo de cálculo se generan nuevas matrices para x_i y y aplicando a cada una logaritmo en base 10. El modelo (*regress*) tiene como función objetivo el valor de la covarianza o coeficiente de determinación R^2 , el cual determina que tan buenas son aproximaciones a los diferentes Q_p^* que resultan del modelo (Steel & Torrie, 1960).

Generar una ecuación para cada uno de los caudales característicos y en cada uno de los clúster requiere previamente, definir cuál es la pareja o el grupo de variables independientes, que combinadas, tengan mayor relación con la variable dependiente, por lo que se realiza una combinatoria entre los diferentes atributos y con cada una de las parejas conformadas se evalúa el valor de R^2 . Cada ecuación seleccionada está conformada por el grupo de variables que mayor R^2 presente, acompañado de sus respectivos exponentes α y coeficiente β .

4.2.5 Resultados

En la Tabla 4-3 se presentan los resultados generales de la regresión lineal múltiple con dos variables independientes para cada uno de los clúster, se muestra también el valor del coeficiente R^2 promedio de los 15 percentiles característicos y una calificación cualitativa a cada uno, se asigna la calificación de *Malo* para grupos con valor de R^2 promedio inferior a 0.3, *Regular* entre 0.3 y 0.4, *Bueno* entre 0.4 y 0.6 y *Muy Bueno* para valores mayores a 0.6. En la Tabla 4-4 se presenta la misma información pero para los resultados de la regresión lineal múltiple calibrada en cada sub región.

Tabla 4-3. Resultados de *RLM* con 2 variables para cada clúster.

Clúster	Validación	Nº Variables	R^2 Promedio	Calificación
1	11077020	2	0.35	Regular
2	23097040	2	0.79	Muy Bueno
3	35017070	2	0.78	Muy Bueno
4	21227010	2	0.23	Malo
5	13017010	2	0.60	Bueno
6	35027020	2	0.26	Malo
7	21207960	2	0.71	Muy Bueno
8	51027020	2	0.40	Bueno
9	42067010	2	0.47	Bueno
10	15017010	2	0.38	Regular
11	32077100	2	0.71	Muy Bueno

Clúster	Validación	N° Variables	R ² Promedio	Calificación
12	23057010	2	0.29	Regular
13	21017020	2	0.23	Malo
14	35027150	2	0.26	Malo
15	21197030	2	0.26	Malo

Tabla 4-4. Resultados de RLM con 2 variables para cada sub región.

Región	Validación	N° Variables	R ² promedio	Calificación
Caribe	13047040	2	0.31	Regular
Magdalena Cauca	21147080	2	0.33	Regular
Orinoquía	35087010	2	0.41	Bueno
Amazonía	44117010	2	0.60	Bueno
Pacífico	52027030	2	0.46	Bueno

Se muestra también un ejemplo (clúster N° 8) de los resultados de cada una de las ecuaciones de regresión obtenidas con dos variables en la Tabla 4-5 para los clúster. En la Tabla 4-6 se muestra el mismo ejemplo pero para la sub región Magdalena Cauca. Ambas tablas muestran la estructura de la ecuación (4-3), las variables independientes x_1 y x_2 se interpretan según la abreviatura indicada en la Tabla 2-1, cabe anotar que en estos resultados ya se muestra la pareja de variables explicativas que mayor R^2 presentan en cada uno de los percentiles, es decir, no hay otra combinación posible, dado ese número de variables, en cada clúster y percentil en específico con la cual se obtengan mejores resultados.

Tabla 4-5. Ejemplo de resultados de RLM - clúster 8.

Caudal Q_p		X_1	α_1	X_2	α_2	R ²
Q_{100}	1.06×10^5	Komp	3.055	%Bosq	0.3464	0.31
Q_{90}	5.66×10^3	Komp	1.763	%Bosq	0.1481	0.34
Q_{80}	2.50×10^2	Komp	1.0989	%Bosq	0.1189	0.38
Q_{70}	6.84×10^2	Komp	0.6775	%Bosq	0.1016	0.44
Q_{60}	3.04×10^3	%Bosq	0.0916	DenD	0.5934	0.51
Q_{50}	1.29×10^2	%Bosq	0.0777	DenD	0.4582	0.50
Q_{40}	7.05×10^2	%Bosq	0.0594	DenD	0.2927	0.54
Q_{35}	8.77×10^1	%Bosq	0.0441	%Urb	0.0099	0.52
Q_{30}	9.75×10^1	%Bosq	0.0326	DenF	0.0147	0.53
Q_{20}	3.96×10^1	DenF	0.0170	Pmáx	0.1164	0.19
Q_{10}	5.03×10^1	%Bosq	-0.0279	Pmáx	0.1647	0.27

Caudal Q_p		X_1	α_1	X_2	α_2	R^2
Q_5	7.39	%Bosq	-0.0377	%Arc	-0.0149	0.27
Q_1	8.43×10^2	%Bosq	-0.0838	Pmed	-0.2748	0.41
$Q_{0.5}$	8.24×10^4	%Bosq	-0.0831	Emed	-0.4112	0.40
$Q_{0.1}$	1.39×10^{14}	Emed	-1.1324	Pmed	-0.7189	0.45

Tabla 4-6. Ejemplo de resultados de RLM - sub región Magdalena Cauca.

Caudal Q_p		X_1	α_1	X_2	α_2	R^2
Q_{100}	1.52×10^{-14}	Emax	1.548	Pmed	3.153	0.28
Q_{90}	8.27×10^{-06}	Emax	0.698	Pmed	0.941	0.24
Q_{80}	7.49×10^{-05}	Emax	0.589	Pmed	0.751	0.39
Q_{70}	3.78×10^{-04}	Emax	0.502	Pmed	0.617	0.43
Q_{60}	1.42×10^{-03}	Emax	0.432	Pmed	0.506	0.45
Q_{50}	4.75×10^{-03}	Emax	0.362	Pmed	0.416	0.43
Q_{40}	2.55×10^{-02}	Emax	0.248	Pmed	0.304	0.37
Q_{35}	5.89×10^{-02}	Emax	0.191	Pmed	0.250	0.32
Q_{30}	4.85×10^{-01}	%Bosq	0.017	Emax	0.095	0.27
Q_{20}	1.340	%Bosq	0.005	DenF	0.008	0.09
Q_{10}	1.14×10^{01}	Emax	-0.116	Pmed	-0.170	0.24
Q_5	6.05×10^{01}	Emax	-0.202	Pmed	-0.313	0.35
Q_1	5.59×10^{01}	ADre	-0.095	Pmed	-0.404	0.34
Q_{100}	9.82×10^{01}	ADre	-0.114	Pmed	-0.457	0.35
Q_{90}	2.26×10^{02}	ADre	-0.155	Pmed	-0.503	0.37

De manera generalizada se obtienen valores de R^2 mayores y por tanto resultados más confiables en el análisis por clúster que en el análisis por sub regiones, esto conlleva a profundizar en el enfoque de agrupación por regiones, que es de por sí uno de los objetivos de este trabajo. Con el fin de verificar si existe alguna variación importante en la calidad de los resultados promedio, entre las diferentes magnitudes de caudales se representa gráficamente en la Figura 4-1, la variación de la covarianza en los diferentes porcentajes de excedencia, se puede apreciar que a diferencia de un valle en el Q_{20} y una disminución del R^2 hacia los caudales mínimos no hay ninguna tendencia significativa.

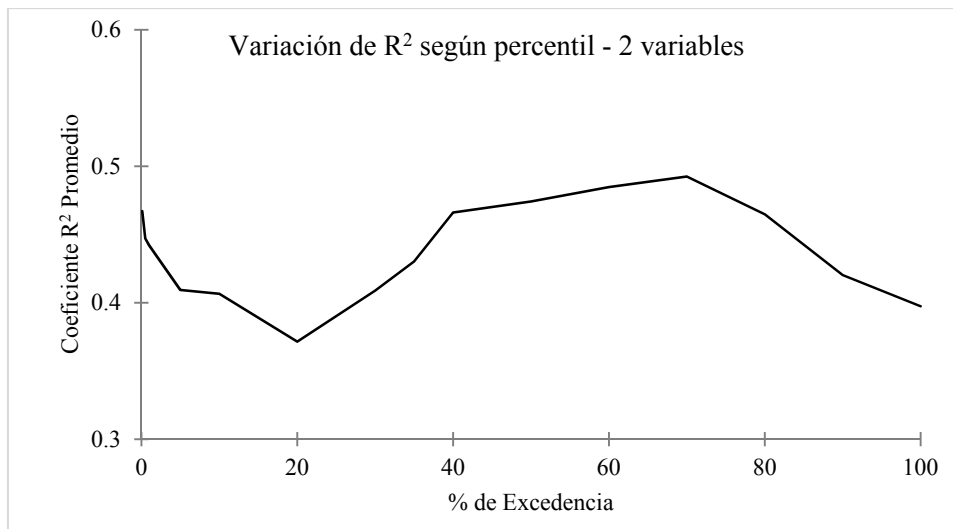


Figura 4-1. Variación del R^2 promedio en los diferentes percentiles de la CDC.

También se puede analizar, cuáles son los atributos más frecuentes en las regresiones de dos variables, se observa el histograma de frecuencias en la Figura 4-2 en el cual se muestran las variables más influyentes en orden descendente. La precipitación media mensual de la cuenca (variable climática), el porcentaje de bosque (variable que depende de la cobertura vegetal), la longitud del cauce principal (variable relacionada con el tamaño de la cuenca) y la pendiente media de la cuenca (variable topográfica), son las más destacadas en este análisis, todas ellas independientes entre sí y con orígenes o características distintas. En contraste atributos como porcentaje de limos y los cuatro percentiles de la curva hipsométrica son los menos frecuentes entre las diferentes ecuaciones. Se tiene en cuenta que este conteo se realizó en los 15 clúster y en los 15 percentiles de cada uno, y dos variables independientes por cada ecuación, por lo tanto se cuenta un total de 450 variables.

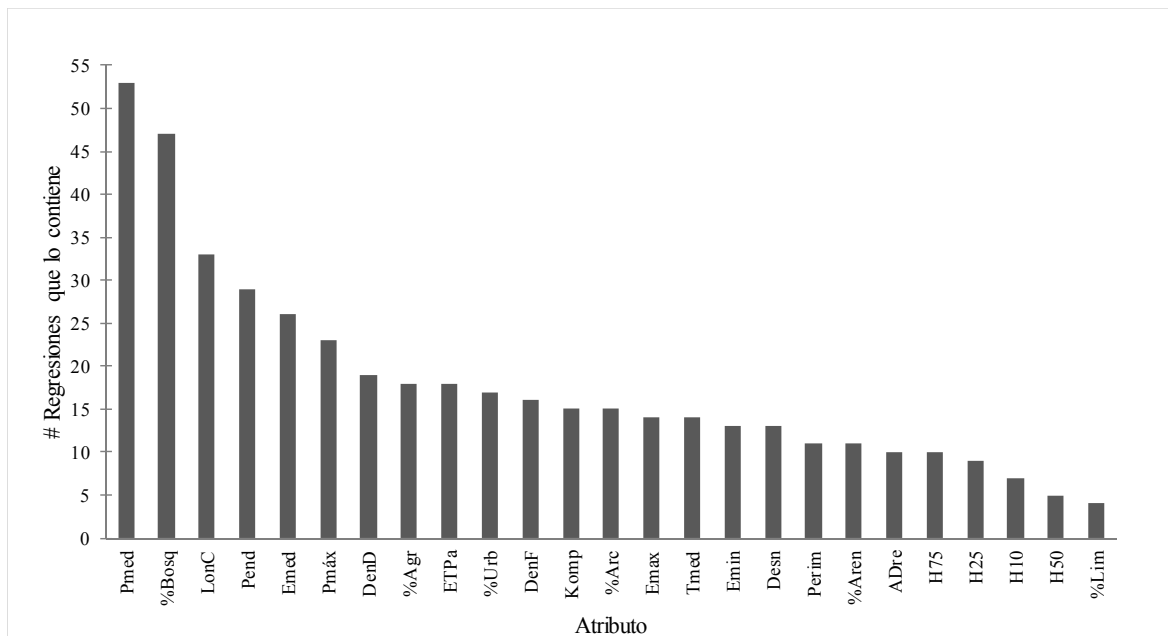


Figura 4-2. Frecuencia de cada atributo en regresiones de dos variables.

Con el fin de mejorar los valores de R^2 en los clúster cuyos resultados preliminares recibieron la calificación de malos o regulares, se recalculan las regresiones múltiples pero esta vez, revisando si al agregar una o varias variables adicionales a cada ecuación de regresión, se observa alguna mejora en los resultados. Se utilizan hasta 5 variables por regresión, incrementando de a una según sea o no satisfactorio el cambio, y dejando los clúster de buena calificación con sus resultados originales. En la Tabla 4-7 se muestra el resumen de resultados de las regresiones con dos o más variables independientes.

Tabla 4-7. Resultados generales de *RLM* en los clúster con 2 o más variables.

Clúster	Validación	R^2 Promedio	Nº Variables	Calificación
1	11077020	0.46	4	Bueno
2	23097040	0.79	2	Muy Bueno
3	35017070	0.78	2	Muy Bueno
4	21227010	0.40	5	Bueno
5	13017010	0.60	2	Bueno
6	35027020	0.53	5	Bueno
7	21207960	0.71	2	Muy Bueno
8	51027020	0.48	4	Bueno
9	42067010	0.54	3	Bueno
10	15017010	0.49	4	Bueno

Clúster	Validación	R ² Promedio	Nº Variables	Calificación
11	32077100	0.71	2	Muy Bueno
12	23057010	0.43	4	Bueno
13	21017020	0.35	5	Regular
14	35027150	0.36	5	Regular
15	21197030	0.49	5	Bueno

Comparando estos nuevos resultados con los presentados anteriormente en la Tabla 4-3, si bien el cálculo se hace más complejo debido al mayor número de variables, puede notarse como en general los resultados mejoran de forma considerable con excepción de dos clúster (13 y 14) que no alcanzaron una mejor calificación que *Regular*. En la Tabla 4-8 se presentan los resultados del clúster N° 8 a manera de ejemplo y como ventana de comparación con los resultados preliminares del mismo grupo mostrados en la Tabla 4-5. En contraste, se pasa de tener un R^2 promedio de 0.40 con dos variables a uno de 0.48 con cuatro.

Tabla 4-8. Ejemplo de resultados de *RLM* clúster N° 8 con 4 variables.

Q _p		X ₁	α_1	X ₂	α_2	X ₃	α_3	X ₄	α_4	R ²
Q ₁₀₀	1.5x10 ⁻⁰³	ADre	-1.055	%Bosq	0.352	DenF	-0.297	LonC	2.491	0.42
Q ₉₀	5.5	Komp	2.044	%Bosq	0.140	H ₁₀	-4.044	H ₂₅	3.195	0.45
Q ₈₀	3.0	Komp	1.289	%Bosq	0.114	H ₁₀	-2.655	H ₂₅	2.070	0.47
Q ₇₀	2.2	Komp	0.812	%Bosq	0.098	H ₁₀	-1.843	H ₂₅	1.420	0.51
Q ₆₀	0.24	%Bosq	0.091	%Urb	0.013	DenD	0.718	Pmáx	-0.231	0.55
Q ₅₀	0.35	%Bosq	0.077	%Urb	0.012	DenD	0.547	Pmáx	-0.172	0.54
Q ₄₀	0.37	%Agr	-0.005	%Bosq	0.057	%Urb	0.011	DenD	0.261	0.57
Q ₃₅	0.93	%Agr	-0.006	%Bosq	0.043	%Urb	0.009	DenF	0.012	0.55
Q ₃₀	0.78	%Agr	-0.005	%Bosq	0.030	DenF	0.020	Emin	0.037	0.58
Q ₂₀	0.95	Komp	-0.234	DenF	0.018	Pend	-0.112	Pmáx	0.140	0.32
Q ₁₀	2.3	%Bosq	-0.026	DenD	-0.345	Pend	-0.134	Pmáx	0.236	0.36
Q ₅	1.5	ADre	0.043	%Bosq	-0.039	%Urb	-0.017	%Arc	-0.026	0.34
Q ₁	5.1 x10 ⁰²	%Bosq	-0.054	%Arc	-0.022	Emed	-0.414	Pmed	-0.344	0.49
Q _{0.5}	0.37	Komp	-0.745	%Bosq	-0.103	H ₁₀	2.569	H ₂₅	-1.897	0.51
Q _{0.1}	5.3 x10 ⁰⁶	%Aren	-0.044	Emed	-1.087	Pend	-0.595	Pmed	-0.650	0.59

Adicionalmente, se presenta el análisis de variación de R^2 entre percentiles actualizado en la Figura 4-3, en la cual se observa que no hay una variación importante en los picos y valles de estos valores en comparación con la Figura 4-1, sin embargo puede observarse que los R^2 incrementan de un gráfico a otro denotando evolución positiva (creciente). En la

Figura 4-4 se muestra el número de regresiones que contiene cada uno de los atributos actualizados a las regresiones con dos o más variables, dentro de los atributos más influyentes se encuentran nuevamente porcentaje de bosque y longitud del cauce principal, pero esta vez, acompañados por porcentaje de suelo agrícola y porcentaje de suelo urbano.

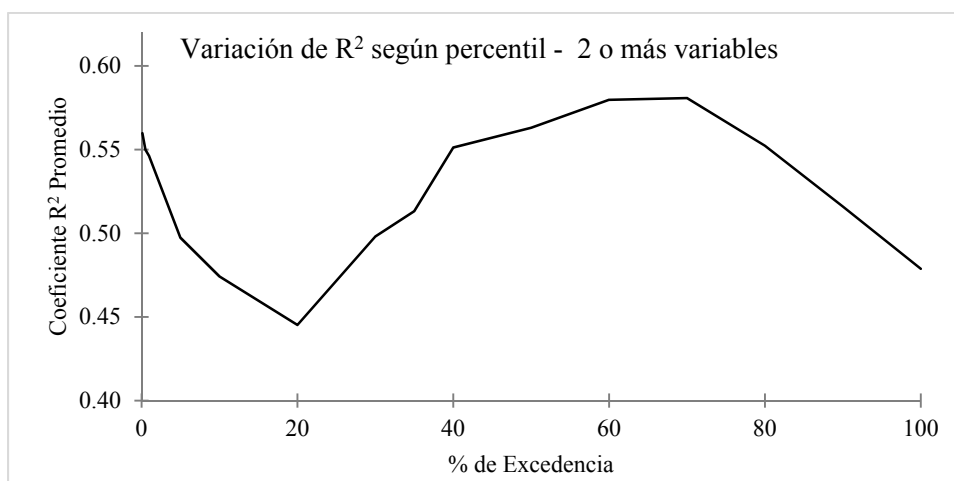


Figura 4-3. Variación de R^2 según percentil en regresiones con 2 o más variables.

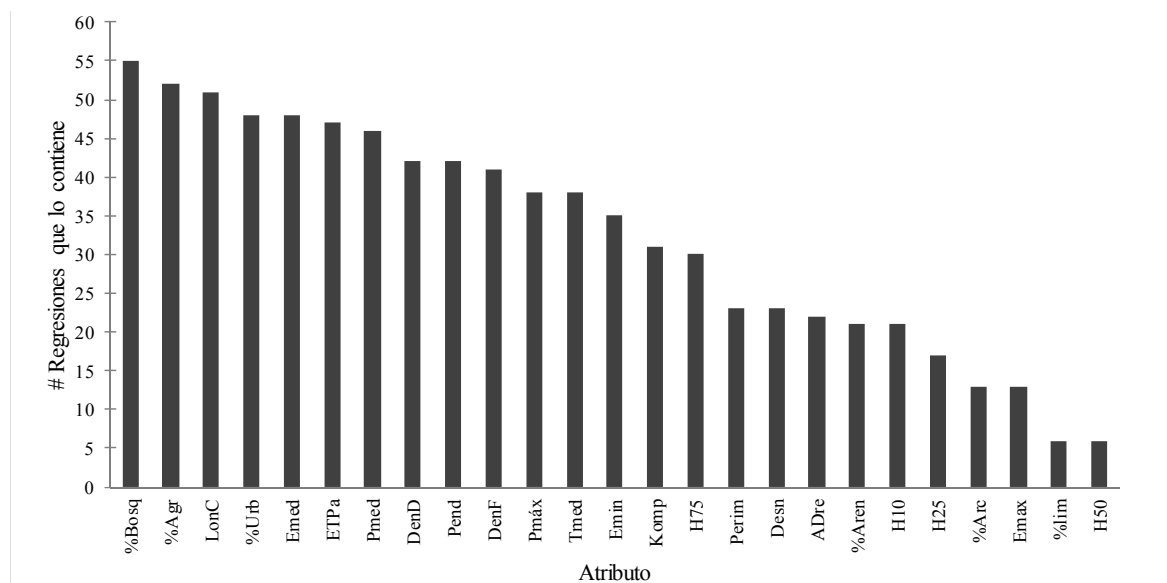


Figura 4-4. Frecuencia de cada atributo en regresiones de dos o más variables.

Se presentan los resultados de ecuaciones de regionalización definitivas para los clúster y para las sub regiones en el Anexo C. Se muestran dos ejemplos de gráficos comparativos

de las CDC estimadas según la metodología de las ecuaciones empíricas regionales (Figura 4-5 y Figura 4-7) con su correspondientes dispersiones de valores estimados vs observados (Figura 4-6 y Figura 4-8). También se presenta un ejemplo de la CDR real y estimada de una estación de validación en la sub región Magdalena Cauca (Figura 4-9) y su respectivo gráfico de dispersión de valores estimados vs observados en la Figura 4-10.

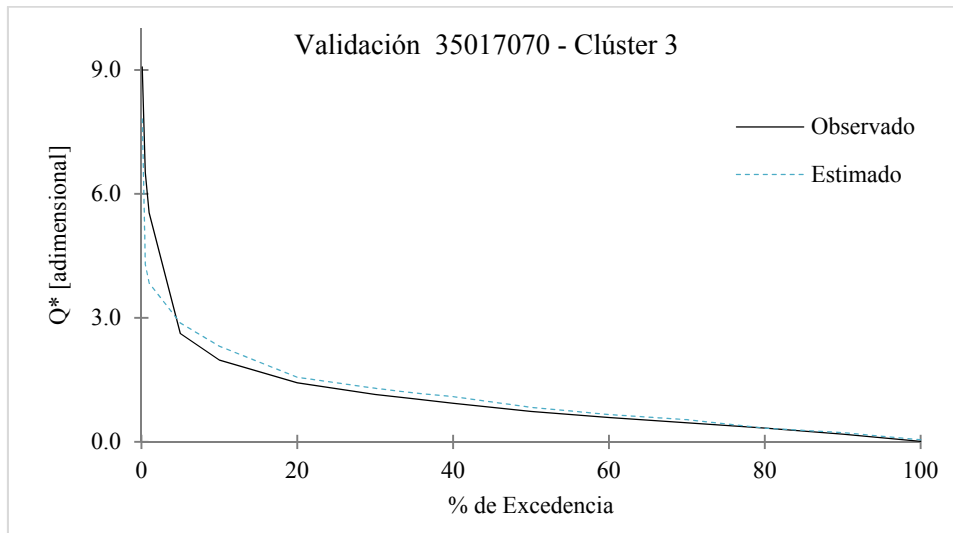


Figura 4-5. Ejemplo de CDC real y estimada, estación 35017070 – Clúster 3, río Orotoy, área de drenaje: 167 km².

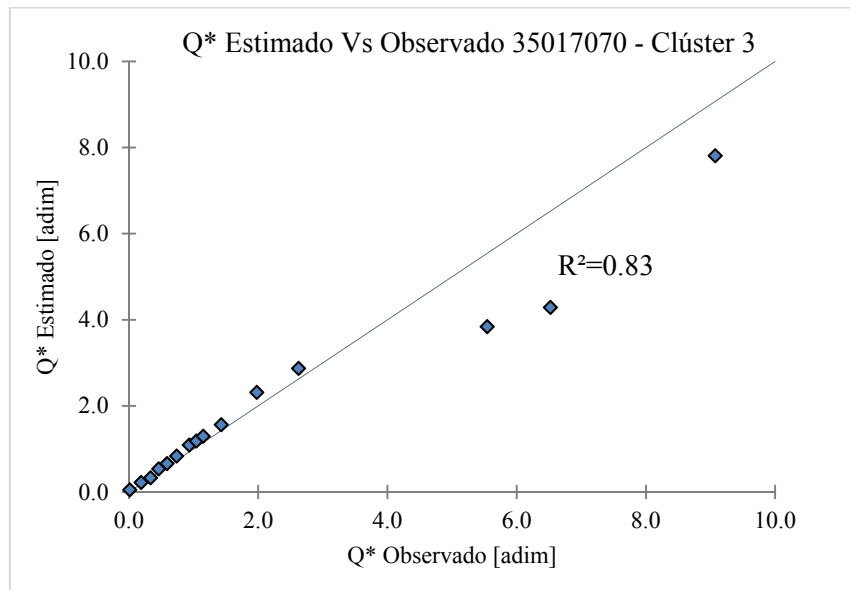


Figura 4-6. Caudal adimensional real vs Caudal adimensional estimado por *RLM* en estación 35017070 y recta $y=x$.

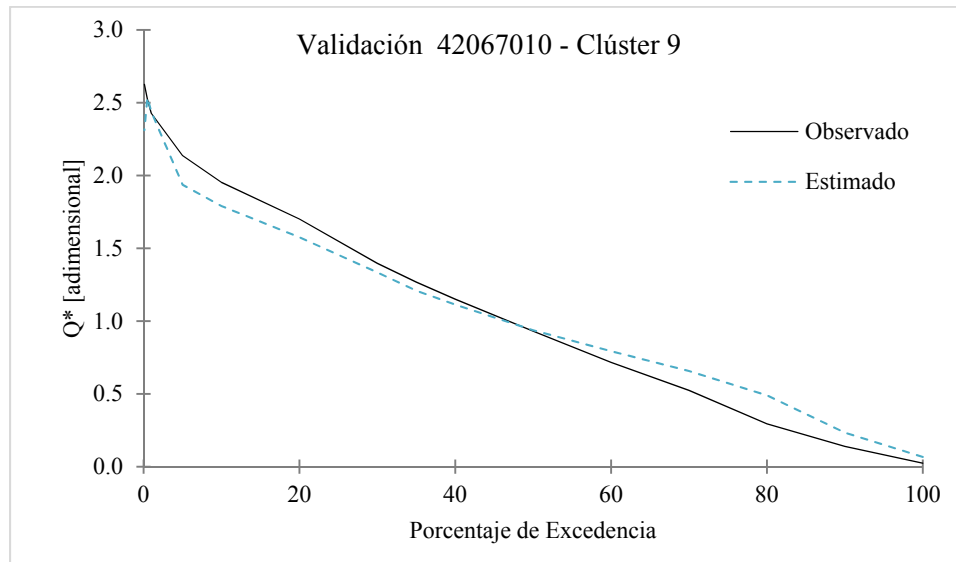


Figura 4-7. Ejemplo de CDC real y estimada, estación 42067010 – Clúster 9, río Vaupés, área de drenaje: 17077 km².

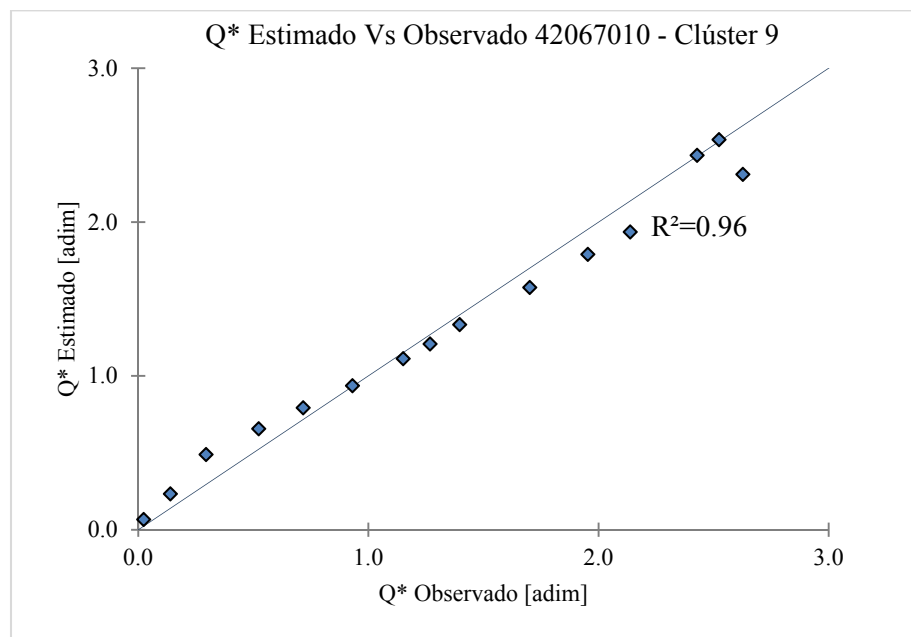


Figura 4-8. Caudal adimensional real vs Caudal adimensional estimado mediante *RLM* en estación 42067010 y recta $y=x$.

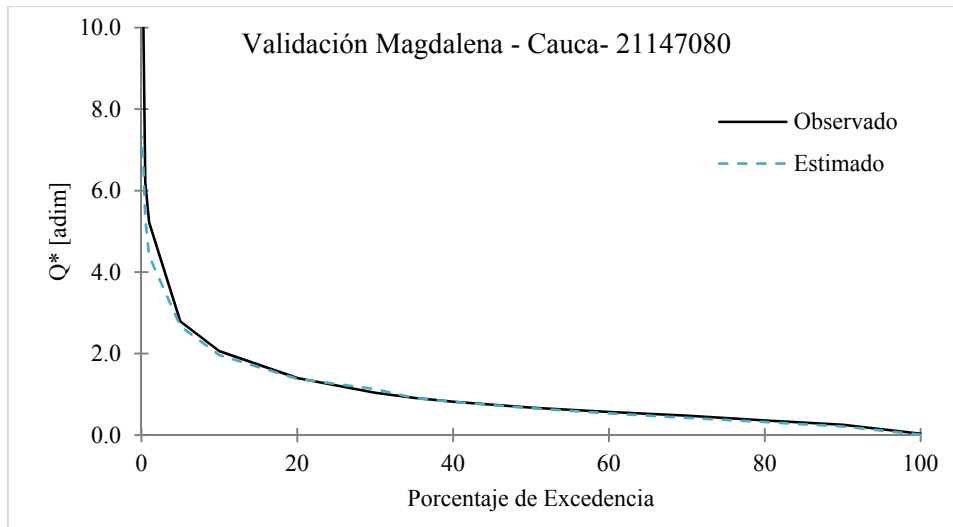


Figura 4-9. Ejemplo de CDC real y estimada, estación 21147080 – sub región Magdalena Cauca, río Cabrera, área de drenaje: 1185 km².

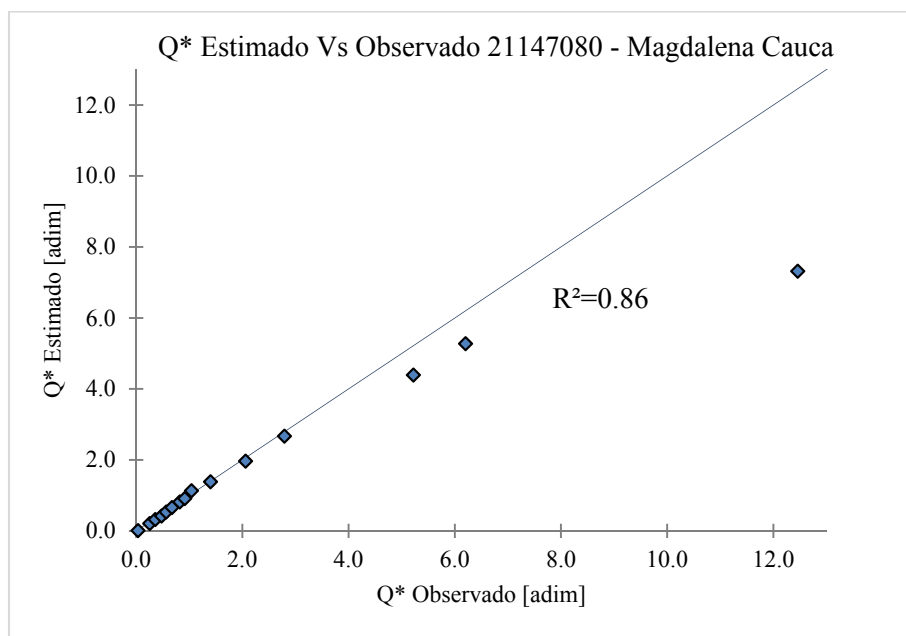


Figura 4-10. Caudal adimensional real vs Caudal adimensional estimado mediante *RLM* en la estación 21147080 y recta $y=x$.

Se evalúa el coeficiente de correlación lineal R y la covarianza R^2 entre los valores estimados - reales vs la recta $y=x$, en la Tabla 4-9 para cada uno de los clúster y en la Tabla 4-10 para cada una de las sub regiones.

Tabla 4-9. Valores de R y R^2 para observados - estimados vs $y=x$ en cada clúster.

Clúster	Estación	$R - y=x$	$R^2 - y=x$
1	11077020	0.86	0.74
2	23097040	0.98	0.97
3	35017070	0.91	0.83
4	21227010	0.79	0.62
5	13017010	0.93	0.86
6	35027020	0.93	0.86
7	21207960	0.80	0.63
8	51027020	0.96	0.92
9	42067010	0.98	0.96
10	15017010	0.82	0.67
11	32077100	0.95	0.90
12	23057010	0.96	0.91
13	21017020	0.91	0.82
14	35027150	0.91	0.82
15	21197030	0.83	0.70

Tabla 4-10. Valores de R y R^2 para estimados - observados vs $y=x$ en cada sub región.

Región	Estación	$R - y=x$	$R^2 - y=x$
Caribe	13047040	0.99	0.98
Magdalena Cauca	21147080	0.86	0.75
Orinoquía	35087010	0.94	0.89
Amazonía	44117010	0.94	0.88
Pacífico	52027030	0.96	0.92

Se calcula el error relativo porcentual, que resulta del valor absoluto de la diferencia entre el caudal observado y el estimado sobre el caudal observado. En la Tabla 4-11 y en la Tabla 4-12 se muestran estos errores calculados para las validaciones correspondientes a los clúster y a las sub regiones respectivamente. En la Figura 4-11 se presenta una comparación entre la variación del error promedio en ambas casos de regionalización, se puede observar tanto en la figura como en las tablas mencionadas, que se obtuvieron mejores aproximaciones en las estimaciones por sub regiones que por clúster.

Tabla 4-11. Errores relativos para validaciones en los clúster.

Clúster	Q ₁₀₀	Q ₉₀	Q ₈₀	Q ₇₀	Q ₆₀	Q ₅₀	Q ₄₀	Q ₃₅	Q ₃₀	Q ₂₀	Q ₁₀	Q ₅	Q ₁	Q _{0.5}	Q _{0.1}
1	36.1	86.5	57.4	48.8	44.5	40.6	27.9	12.2	10.3	25.5	16.9	2.0	39.5	54.4	112.1
2	143.0	0.4	1.9	3.5	0.1	2.1	2.9	2.2	5.8	4.0	1.2	2.9	7.1	7.3	11.7
3	368.4	20.2	0.3	17.1	12.7	13.7	17.4	13.8	13.2	9.6	17.1	9.6	30.7	34.2	13.9
4	99.7	71.4	62.9	65.0	60.7	29.5	16.7	11.6	17.2	39.7	15.5	16.9	43.4	44.2	57.2
5	36.5	10.0	9.7	14.9	15.7	11.7	5.9	2.9	8.3	3.5	13.3	17.9	47.0	45.4	55.7
6	70.7	18.9	18.5	6.1	3.3	4.5	5.8	3.0	7.0	11.2	5.4	0.5	18.6	18.5	34.1
7	33.0	67.5	62.7	60.0	57.3	19.6	4.4	24.4	28.2	30.8	22.2	37.6	128.2	130.7	162.1
8	123.8	6.7	0.8	1.6	8.1	9.3	8.2	9.5	4.8	1.6	4.6	13.3	7.5	20.2	22.2
9	188.0	67.7	66.5	25.6	10.8	0.8	3.3	4.6	4.4	7.3	8.3	9.4	0.3	0.5	12.0
10	96.6	68.8	22.3	15.5	21.9	30.4	37.3	32.7	32.6	9.6	2.1	49.0	66.1	12.8	1.0
11	43.6	18.1	23.5	29.4	20.5	2.6	2.3	7.7	8.7	8.4	15.6	8.6	22.4	2.0	0.4
12	59.3	34.8	0.6	6.0	0.5	9.2	6.8	4.9	4.4	3.7	8.1	7.3	12.9	9.1	12.2
13	85.1	17.0	21.5	8.1	1.9	4.0	0.6	2.7	4.4	5.3	10.8	5.7	0.1	2.6	22.7
14	76.4	4.3	8.7	4.7	22.5	10.2	3.9	2.1	0.4	7.6	4.4	100.0	47.9	37.5	66.6
15	1475.6	157.1	81.0	59.2	44.7	34.3	6.4	8.6	8.3	5.8	19.4	27.9	35.9	41.9	54.1
media	195.7	43.3	29.2	24.3	21.7	14.8	10.0	9.5	10.5	11.6	11.0	20.6	33.8	30.8	42.5

Tabla 4-12. Errores relativos para validaciones en sub regiones.

Sub Región	Q ₁₀₀	Q ₉₀	Q ₈₀	Q ₇₀	Q ₆₀	Q ₅₀	Q ₄₀	Q ₃₅	Q ₃₀	Q ₂₀	Q ₁₀	Q ₅	Q ₁	Q _{0.5}	Q _{0.1}
Caribe	17.86	29.6	13	5.4	4.1	7.7	9	6.6	21	13	2.5	9.18	19.6	3.94	7.9
Magdalena	64.41	16.9	11	11	5.4	1.7	0.3	0.3	8.9	1.2	4.5	4.25	15.9	14.9	41.3
Orinoquía	34.44	13.4	34	33	25	17	12	3.8	1.1	3.1	9.9	15.5	15.5	19.4	1.7
Amazonía	59.35	35.3	14	16	7.9	7.4	4.6	2.3	11	3.2	16	7.39	43.4	54.8	69.5
Pacífico	417.5	15.1	12	8.8	7	6.3	4.5	4.2	3.8	2.7	8.4	14.1	16.6	15	10.9
media:	118.7	22.1	17	15	9.9	8	6	3.4	9.1	4.7	8.3	10.1	22.2	21.6	26.2

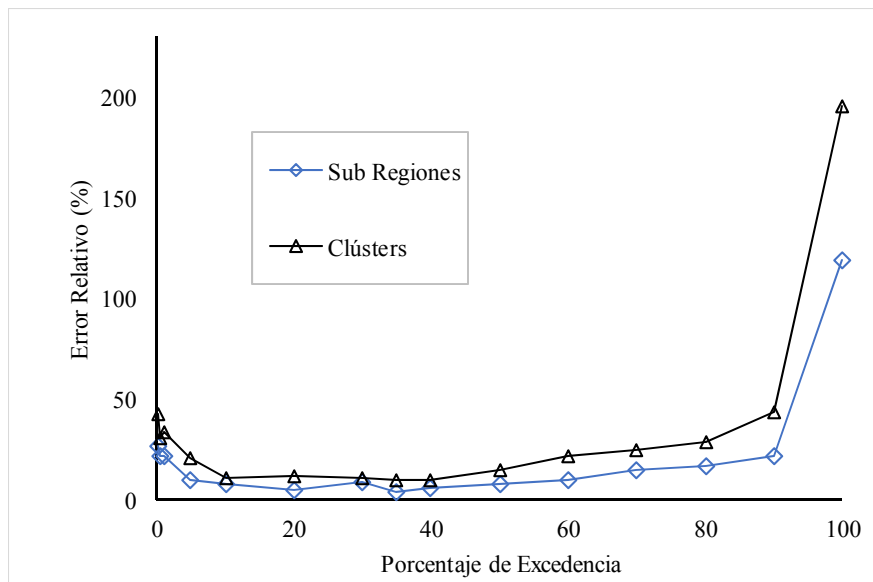


Figura 4-11. Comparación del error relativo entre la estimación mediante regionalización por clúster y por sub regiones.

4.3 Análisis de resultados

Como resultado principal, se tiene una ecuación empírica que permite estimar cada uno de los percentiles característicos de caudal adimensional Q_p^* de cada una de las regiones hidrológicas de Colombia definidas en el capítulo 3 y para cada una de las sub regiones del país predefinidas por el IDEAM. De manera general, se tienen mejores correlaciones en las ecuaciones de los 15 clúster en relación con las ecuaciones de las 5 sub regiones, por lo tanto se recomienda a priori aplicar el modelo empleando la regionalización acorde a las regiones delimitadas en este trabajo, sin embargo, la ubicación de una cuenca objeto de la estimación, se hace más simple mediante el análisis por sub regiones.

En los casos de validación de las ecuaciones de regresión, se obtiene un valor promedio de error relativo de aproximadamente 27%, repartido este en aproximadamente 50% en la aproximación a caudales mínimos, 9% para caudales medios y 26% para caudales máximos. Por otra parte, las validaciones realizadas para las sub cuencas presentaron un mejor desempeño de forma general, observando la Figura 4-11 puede notarse que la línea correspondiente a los clúster señala en casi todas las magnitudes, mayores errores, esto es

diferente de lo que se esperaba, pues las ecuaciones de regresión presentan generalmente mayores valores de R^2 en las ecuaciones calibradas para los clúster.

El error en las aproximaciones varía espacialmente entre regiones hidrológicas, en algunos casos es evidente que se obtienen mejores resultados en las estimaciones de los diferentes Q_p^* , por ejemplo, los clúster 2, 3, 7 y 11 que presentan de manera considerable valores mayores de R^2 respecto a los demás, o los clúster 13 y 14 que en contraste, presentan notablemente los valores más bajos de este coeficiente. Esto no tiene que ver con el número de puntos que se hayan utilizado para la calibración ni con qué tan homogéneas sean las cuencas de un mismo grupo, esto puede revisarse en conjunto con el mapa de distancias euclidianas mostrado en la Figura 3-18.

Se observa también, que aunque los valores de R^2 sean bajos, es posible tener aproximaciones muy buenas en las comparaciones de valores estimados con observados, esto depende de cual sea la corriente y en qué punto de interés se realice la estimación. Ejemplos de esto son las estaciones validadas para los clúster: 8, 9, 10 y 11 en los cuales se presentaron resultados muy buenos pero sin tener, comparativamente, altos valores de coeficientes R^2 , caso contrario ocurre en los clúster 1, 6 y 7, en los cuales, teniendo valores de R^2 entre buenos y muy buenos, los resultados no fueron los esperados.

Las variables tipo cobertura de suelo son notablemente frecuentes en las ecuaciones de regresión múltiple, seguido por las variables climáticas dependientes de la precipitación y la evapotranspiración potencial y luego por las variables que dependen de la topografía de la cuenca como elevaciones, desnivel y pendiente media. Esto puede ayudar a inferir cuáles son los procesos más relevantes a la hora de estimar el régimen de caudales de una cuenca.

En las validaciones de los resultados, puede notarse que los caudales medios (percentiles 30, 35 y 40 de la CDC) son los que mejores aproximaciones alcanzan, los mínimos y los máximos es donde más incertidumbre se tiene. Una posible explicación a esto, es que las estaciones de medición se calibran generalmente durante descargas medias, por lo tanto las mediciones extremas son en su gran mayoría (o su totalidad) extrapolaciones. Esto representa un ruido o una incertidumbre adicional para la calibración y posterior estimación en estas frecuencias, sin embargo, el modelo no parece mostrar variaciones importantes en su desempeño entre los diferentes percentiles si se observa solo los valores promedio de R^2 , con la única excepción del percentil 20 el cual corresponde a una transición entre medios y máximos (Figura 4-3).

Se propone que el número máximo de variables que intervienen en una ecuación de regresión es de 5, debido a que en el momento de aplicar el modelo, se haría muy compleja la búsqueda de muchas variables, esto requeriría de más trabajo de manejo de datos en sistemas de información geográfica y como se puede observar, en ocasiones no es muy significativa la mejora por agregar una variable más a cada ecuación.

La forma de la CDC adimensional está fuertemente relacionada con el tamaño de la cuenca, es evidente que en cuencas pequeñas (orden entre 10^1 y 10^2 km²) ésta toma una forma similar a la de una “L” (concavidad pronunciada), es decir hay mucha diferencia entre las magnitudes de los caudales máximos y los medios, mientras que para cuencas de gran tamaño (orden de magnitud de entre 10^4 y 10^5 km²) la forma de la CDC es más plana y suavizada. En las CDC observadas y estimadas mediante el modelo de regresión, puede notarse que el modelo en todos los casos simula bien la forma general de la CDC, esto puede deberse a que muchos atributos ya están de antemano, relacionados directa o indirectamente con el tamaño de la cuenca.

Aunque no fue común dentro de los resultados, es posible que al estimar caudales con esta metodología se presenten errores de tipo conceptual en la reconstrucción de la curva sintética, por definición es imposible o incorrecto que un caudal Q_{p1} con porcentaje de excedencia mayor a otro Q_{p2} ($p1 > p2$) sea mayor, es decir Q_{p1} siempre debe ser menor a Q_{p2} . En caso de ocurrir esta errata en el proceso de estimación, se sugiere realizar una eliminación del percentil en cuestión y una posterior interpolación lineal para recuperar la información descartada.

Se obtuvieron buenos valores de correlación entre las matrices real-estimado y $y=x$ tanto para los clúster como para las sub regiones (Tabla 4-9), esto indica que las estimaciones de caudal adimensional se mantienen dentro de un orden de magnitud coherente, en general las aproximaciones resultaron satisfactorias a juicio del autor.

5 Análisis No Estacionario de la Curva de Duración de Caudales en Colombia

5.1 Introducción y metodología

Se trata de revisar si existen cambios significativos en los diferentes percentiles de la CDC a lo largo del tiempo para diferentes ríos de Colombia, para ello, se cuenta con datos de caudales a escala diaria, de los cuales puede seleccionarse un conjunto de estaciones que cuenten con registros distribuidos temporalmente de tal manera que se abarque un período suficiente de caracterización climática. El período puede definirse de 30 años según la Organización Meteorológica Mundial, WMO por sus siglas en inglés. Si bien en este trabajo no se espera explicar los porqués o las razones de los hipotéticos cambios encontrados, se apunta hacia la realización de aportes en la identificación de cuáles son las cuencas o las regiones y en qué frecuencias varían con mayor impacto las CDC a largo plazo.

La prueba no paramétrica de Mann-Kendall se utiliza como verificación de la existencia de tendencias a largo plazo en cada una de las series seleccionadas para el estudio, esta prueba se lleva a cabo para la realización de este trabajo pues admite series con períodos temporales sin información y porque no necesita que los datos tengan alguna distribución en particular (Gilbert, 1987). Para la realización de la prueba, los datos se deben encontrar en el orden temporal en el que fueron tomados, posteriormente, se puede proceder al cálculo el estadístico S , el cual indica el signo global de la serie, el estadístico está dado según la ecuación (5-1):

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{sgn}(x_j - x_k) \quad (5-1)$$

donde sgn es el signo de la operación y toma valores de +1 y -1 para signos positivo y negativo respectivamente y x_k todo registro inmediatamente anterior a x_j . Se tiene que si el valor de S es considerable y positivo, se sospecha que la serie es creciente en el tiempo, y

de manera análoga, indica que la serie es significativamente decreciente a largo plazo si se tienen grandes magnitudes de S con signo negativo. El estadístico Z por su parte, relaciona el valor de S con la longitud de la serie o el número de registros. La varianza dada según la ecuación (5-2), sirve como insumo para el cálculo de Z tal y como se muestra en las ecuaciones (5-3) y (5-4).

$$Var(S) = \frac{1}{18} \left(n(n-1)(2n+5) - \sum_{i=1}^m t_i(i-1)(2i+5) \right) \quad (5-2)$$

$$Z = \frac{S-1}{\sqrt{Var(S)}} \quad \text{para } S > 0 \quad (5-3) \quad Z = \frac{S+1}{\sqrt{Var(S)}} \quad \text{para } S < 0 \quad (5-4)$$

Debido a que la existencia de autocorrelación en algunas series incrementa la posibilidad de detectar tendencias con la prueba de Mann Kendall, se propone otra metodología para el cálculo de la varianza que se conoce como prueba de Mann Kendall modificada (Hamed & Ramachandra, 1998). Con esta, se busca eliminar los efectos de la autocorrelación presente en algunas series temporales, el cálculo de la varianza en este caso, similar al mostrado en la ecuación (5-2), pero con un término adicional, tal y como se muestra en las ecuaciones (5-5) y (5-6).

$$V^*(S) = Var(S) \frac{n}{n^*} \quad (5-5)$$

$$\frac{n}{n^*} = 1 + \frac{2}{n(n-1)(n-2)} \sum_{i=1}^{n-1} (n-i)(n-i-1)(n-i-2) \rho_s(i) \quad (5-6)$$

Donde $\rho_s(i)$ es la función de autocorrelación de los rangos en las series temporales. Para evaluar si el resultado es estadísticamente significativo, se calcula la probabilidad con una función de distribución normal, el valor de H indica si la hipótesis nula de la prueba es rechazada (hay posibilidad de que exista tendencia, $H=1$) si el valor absoluto de Z es mayor a $Z_{1-\alpha/2}$, mientras que no puede rechazarse ($H=0$) en caso contrario. α corresponde a una confiabilidad a escoger, en este caso es del 95% ($\alpha=0.05$). Con el fin de calcular la magnitud de las diferentes tendencias, se propone un ajuste lineal a toda la serie de datos de la forma:

$$y = mx + b \quad (5-7)$$

$$m = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2} \quad (5-8)$$

donde la variable independiente x es el número de días de la serie temporal y la variable dependiente es el valor del caudal correspondiente en m^3/s . Se pretende obtener el valor del coeficiente m en cada serie seleccionada para evaluar la tendencia y comparar. Se seleccionan dos series de cada sub región hidrológica, una con la mayor tendencia creciente y la otra con la mayor tendencia decreciente, obteniendo 10 series de caudales en total, para comparar en cada una de éstas la CDC construida de forma empírica para diferentes períodos de tiempo. Además se calculan, año a año, caudales característicos con el fin de revisar tendencias entre diferentes magnitudes de caudales, los percentiles seleccionados son: Q_{95} (mínimos), Q_{75} (medios mínimos), Q_{25} (medios máximos) y Q_1 (máximos).

De manera adicional, se calcula el estimador de pendiente de Sen (Sen, 1968) el cual se utiliza en series donde se asume que existe tendencia lineal y se tengan discontinuidades. La estimación de la pendiente está dada según la ecuación (5-9).

$$P_i = \frac{x_i - x_k}{j - k} ; i = 1, 2, \dots, N \quad (5-9)$$

donde x_i y x_k son los datos en los tiempos j y k respectivamente ($j > k$). Para una serie con número de registros igual a n , se tendrán tantas pendientes P_i estimadas como: $N = n(n - 1)/2$, el valor del estimador de pendiente de Sen en cada serie es igual a la mediana de los N valores. Finalmente, para complementar el análisis no estacionario de las series seleccionadas, se estiman tendencias lineales en la varianza mediante el cálculo de la desviación estándar en una ventana móvil de 10 años con paso anual. La desviación estándar está dada según la ecuación (5-10):

$$Std = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}} \quad (5-10)$$

Donde Std es la desviación estándar en m^3/s , \bar{x} y N el promedio y el número de datos de cada ventana móvil respectivamente.

5.2 Selección de series temporales

Se cuenta con la misma base de 655 estaciones de registros de caudales diarios desarrollada en el numeral 2.2.1, de las cuales se seleccionan únicamente estaciones con más de treinta años netos de registros, es decir, que cuenten con por lo menos 10.950 mediciones. La Figura 5-1 muestra la distribución espacial de las estaciones que cumplen con este requisito.

Del análisis son apartadas estaciones que se corresponden o se ubican cerca de tramos de corrientes con múltiples canales, esto pues miden solamente un brazo y es posible que en ellas ocurran cambios en la sección que provoquen el aumento o la disminución del caudal sin que sea necesariamente por cambios en las condiciones o atributos de la cuenca. En el Río Magdalena, a la altura de La Mojana, presenta estas características, la Figura 5-2 muestra la serie temporal de una de estas estaciones a manera de ejemplo.

También se excluyen series temporales de caudal que presenten visiblemente inconsistencias en sus registros tales como como cambios en el método de medición, se muestra un ejemplo de esto en la Figura 5-3 en la cual se observa cómo cambia el valor de los registros en el río Putumayo, conforme se re instala la estación de medición.

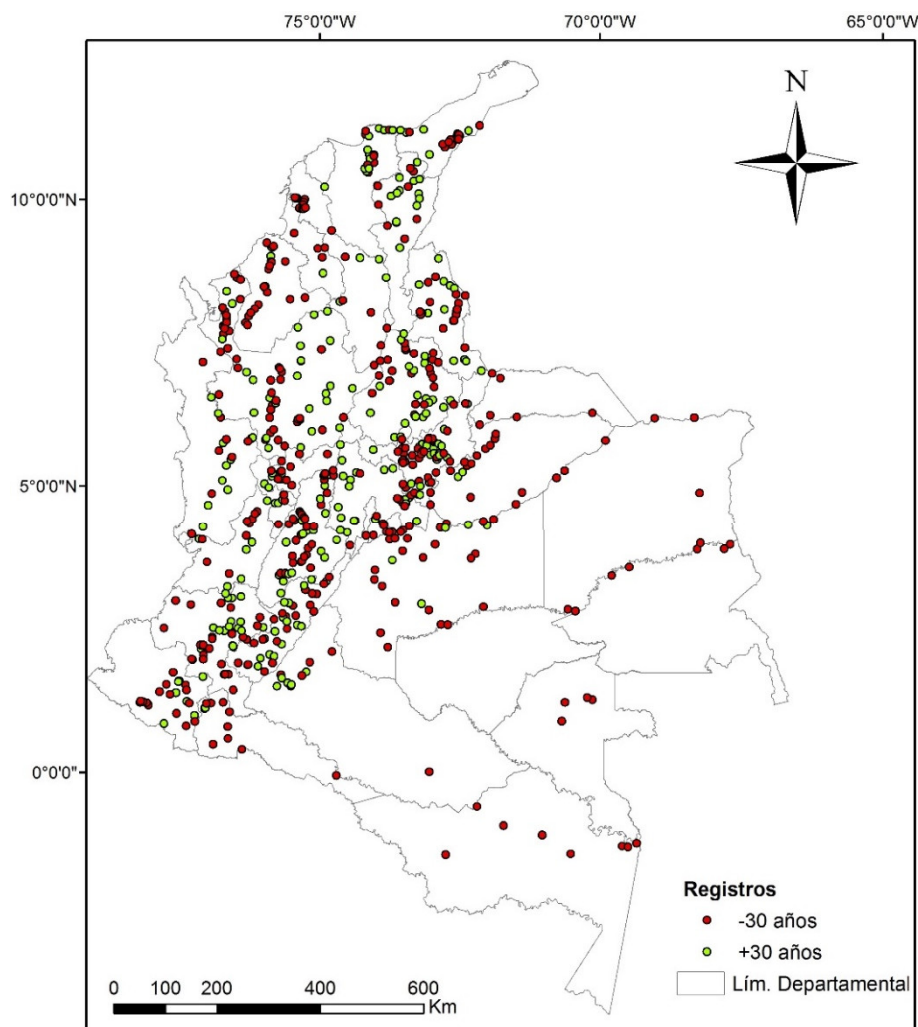


Figura 5-1. Distribución espacial de estaciones de caudal con más de 30 años netos de registros.

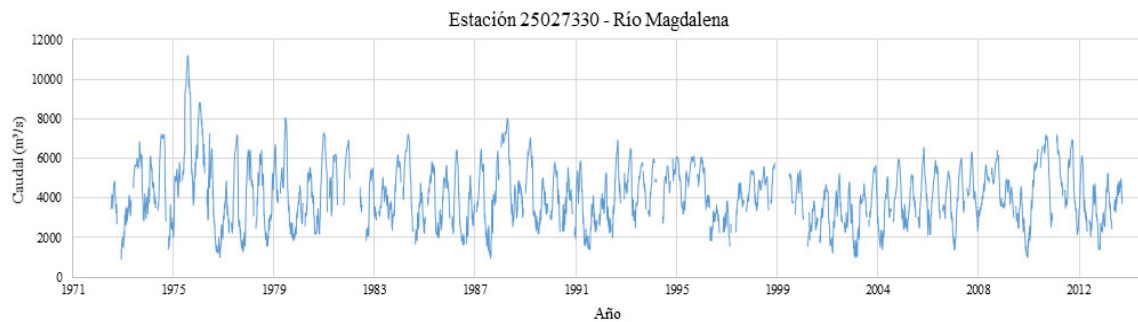


Figura 5-2. Ejemplo de serie temporal descartada por ubicarse en lecho de múltiples canales.

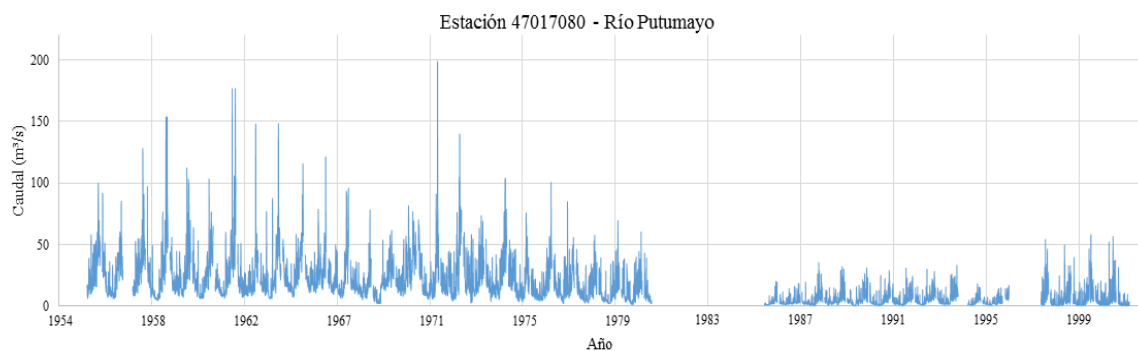


Figura 5-3. Ejemplo de serie descartada por inconsistencias en las mediciones

Se aplica la prueba no paramétrica de Mann-Kendall a cada una de las 655 series temporales de caudal diario y se destaca el estadístico H el cual toma valor de 0 cuando no se puede rechazar la hipótesis nula o de 1 cuando la rechaza. En la Figura 5-4 se ubican espacialmente las estaciones de medición con sus correspondientes valores de H , en adelante, el análisis no estacionario se realizará únicamente para las series en las que éste es igual a 1, pues se descartan las series en las cuales no se puede rechazar la hipótesis nula dada la confiabilidad del 95% ($\alpha=0.05$). Se realiza un filtro de las estaciones que cumplen con los tres requisitos mencionados, más de treinta años de registros netos, ninguna sospecha de inconsistencia en sus mediciones y estadístico $H=1$. En la Figura 5-5 se muestra la localización de las estaciones resultantes después del proceso de filtrado. 222 estaciones en total cumplen con los requisitos.

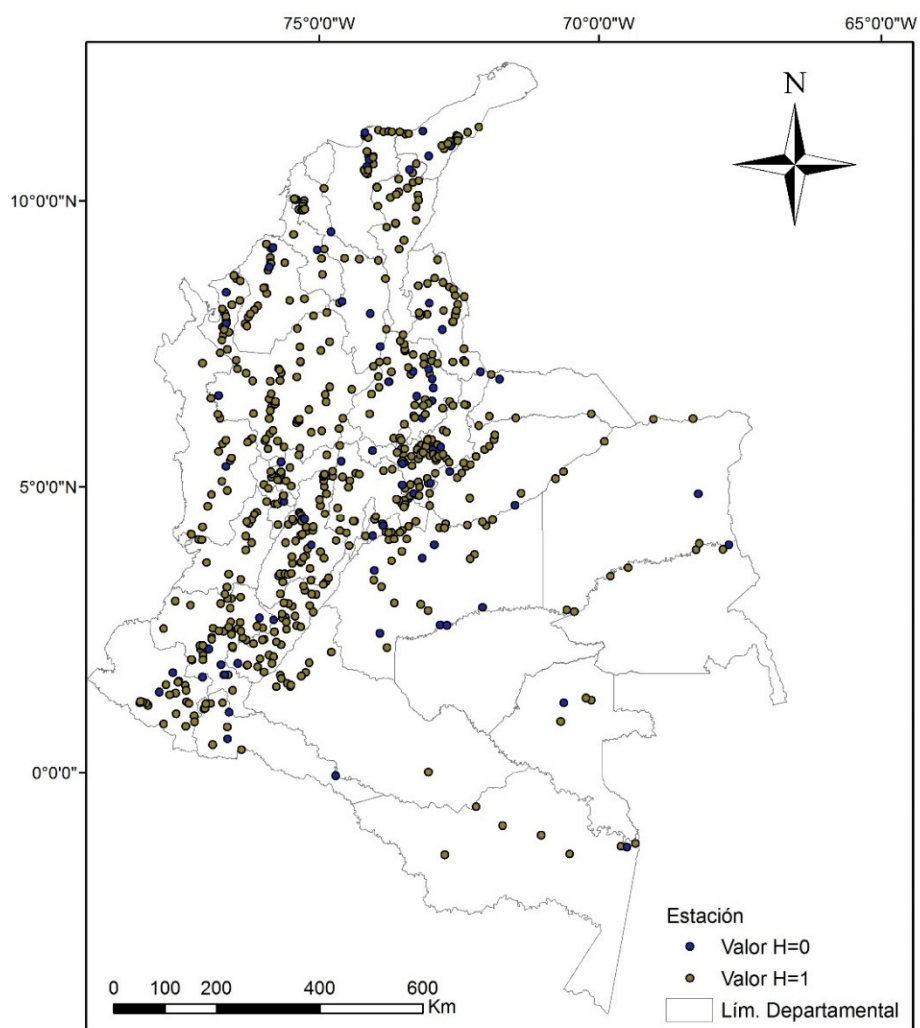


Figura 5-4. Distribución espacial del valor de la prueba H en estaciones de caudal.

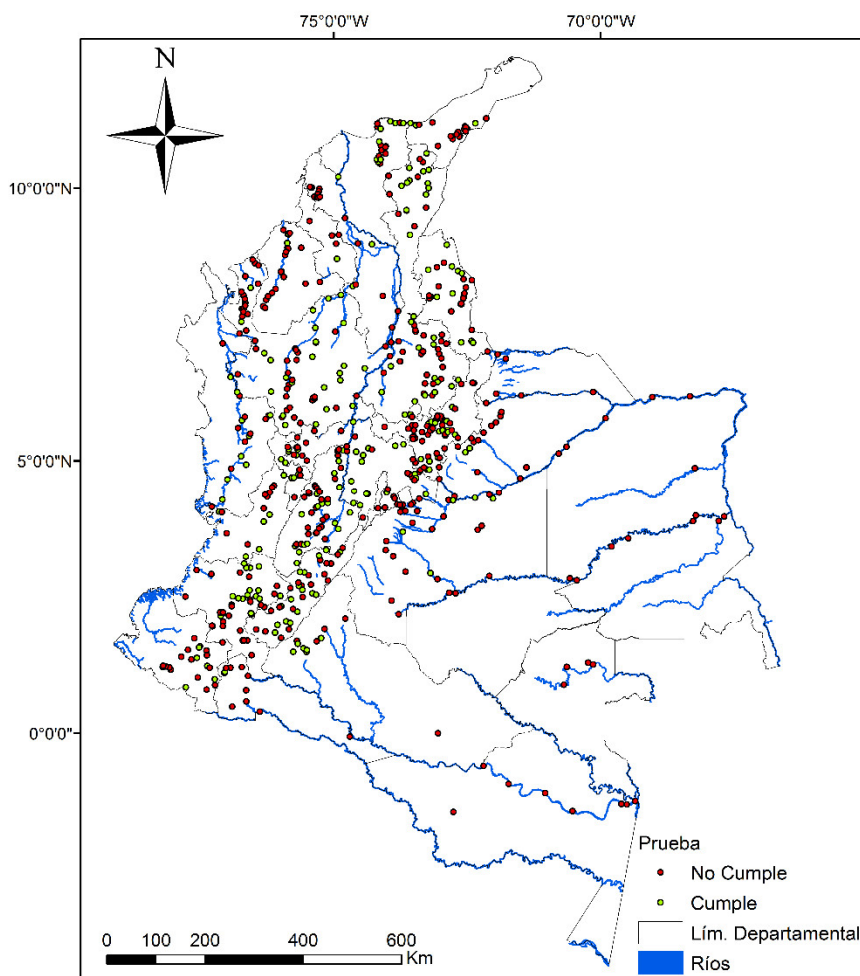


Figura 5-5. Estaciones disponibles para el análisis no estacionario de la CDC.

5.3 Resultados

En la Figura 5-6 se muestra el esquema correspondiente a los valores de las tendencias en cada una de las estaciones dónde éstos fueron calculados, se observa que en general, los caudales tienden a incrementar sus valores en un horizonte de largo plazo en casi todo el territorio nacional, las tendencias son predominantemente crecientes. Las regiones del Chocó y del bajo cauca presentan de forma clara un predominio de series temporales con tendencia creciente, mientras que en ninguna región en el país se observa una mayoría o un patrón de series temporales decrecientes.

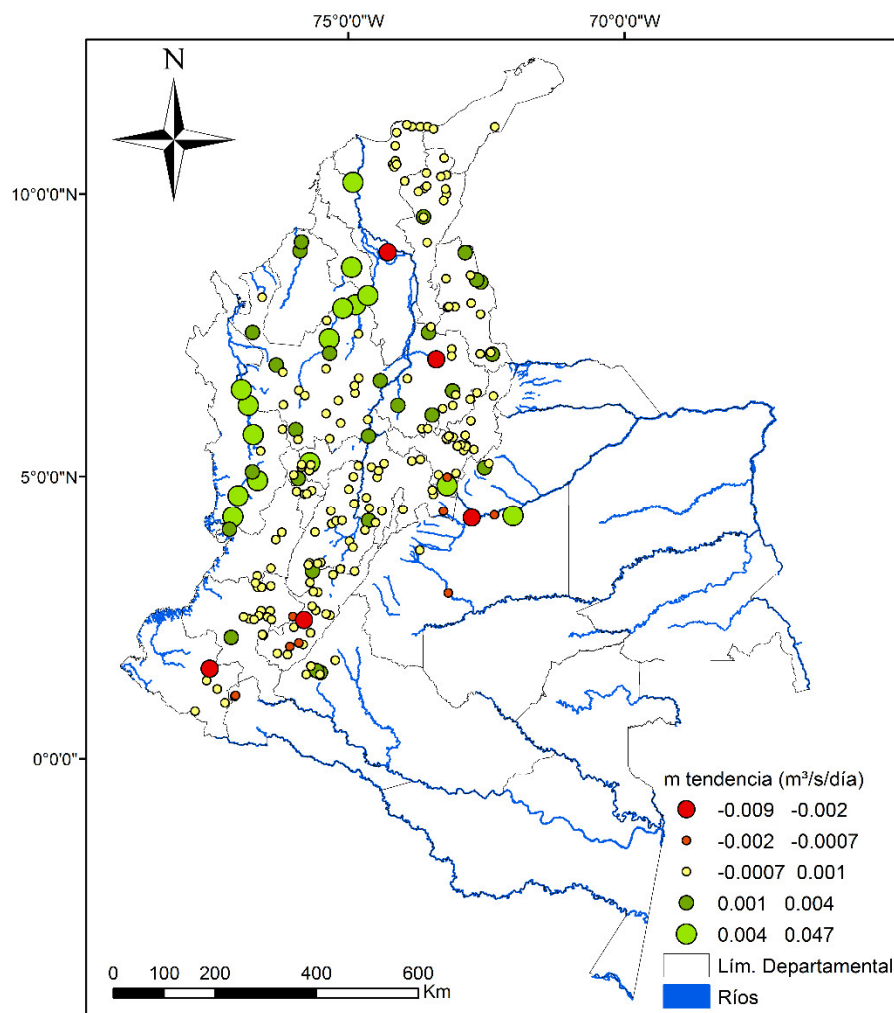


Figura 5-6. Rengos de valores de pendiente de la recta de ajuste en estaciones analizadas.

Los valores obtenidos del estimador de pendiente de Sen presentan una estrecha relación ($R^2=0.87$) con respecto a la pendiente de la recta de ajuste, en la Figura 5-7 se muestra la dispersión de estas dos variables tomando los resultados de las 655 estaciones.

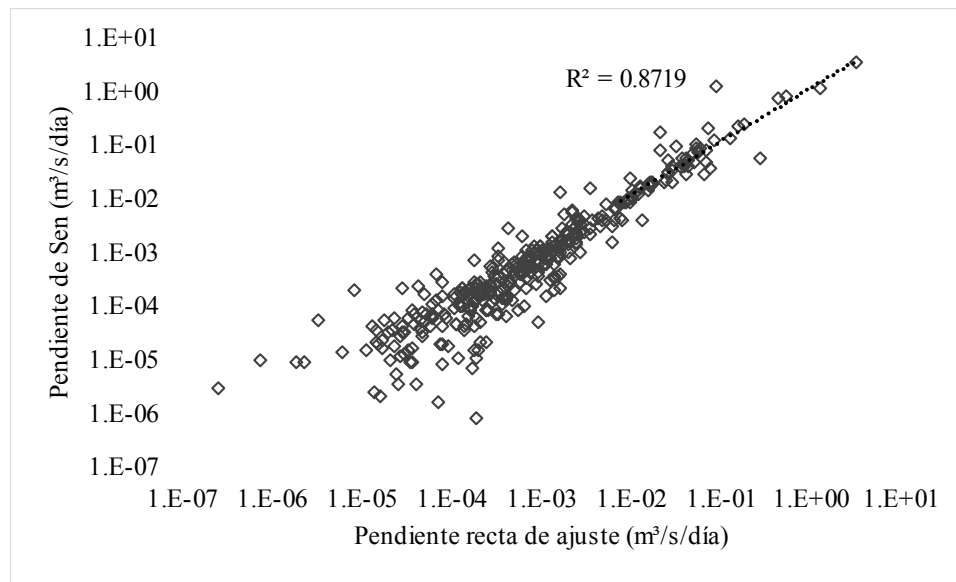


Figura 5-7. Dispersión pendiente de Sen vs pendiente recta de ajuste.

De cada sub región de Colombia se seleccionan tanto la serie con el mayor valor de tendencia creciente como la serie con el mayor valor de tendencia decreciente y se toman diferentes períodos de la CDC de cada una. Para tratar de observar posibles evoluciones en el tiempo, se divide el período de registro en tres intervalos de igual número de años, y se ilustra la CDC de cada uno para contrastar resultados. Entre la Figura 5-8 y la Figura 5-27 se presentan los resultados del análisis no estacionario para las 10 series con su respectiva línea de ajuste. Las CDC se muestran también con el eje de las abscisas en escala logarítmica base 10 para que se puedan distinguir visualmente diferencias entre los caudales máximos ($Q_{0.01}$, $Q_{0.03}$, $Q_{0.1}$, $Q_{0.3}$ y Q_1). En las CDC mostradas a la izquierda con el eje graduado en escala lineal se pueden apreciar las diferencias en los caudales máximos y mínimos, la pendiente de la recta de ajuste se muestra en las series temporales con la letra m. Se muestra además la evolución anual y la tendencia de ajuste lineal, de los caudales característicos correspondientes a los porcentajes de excedencia 95, 75, 25 y 1 en cada una de los ríos seleccionados. Se muestra también la pendiente de Sen (m Sen) y la pendiente del ajuste de la desviación estándar vs el año en cada serie.

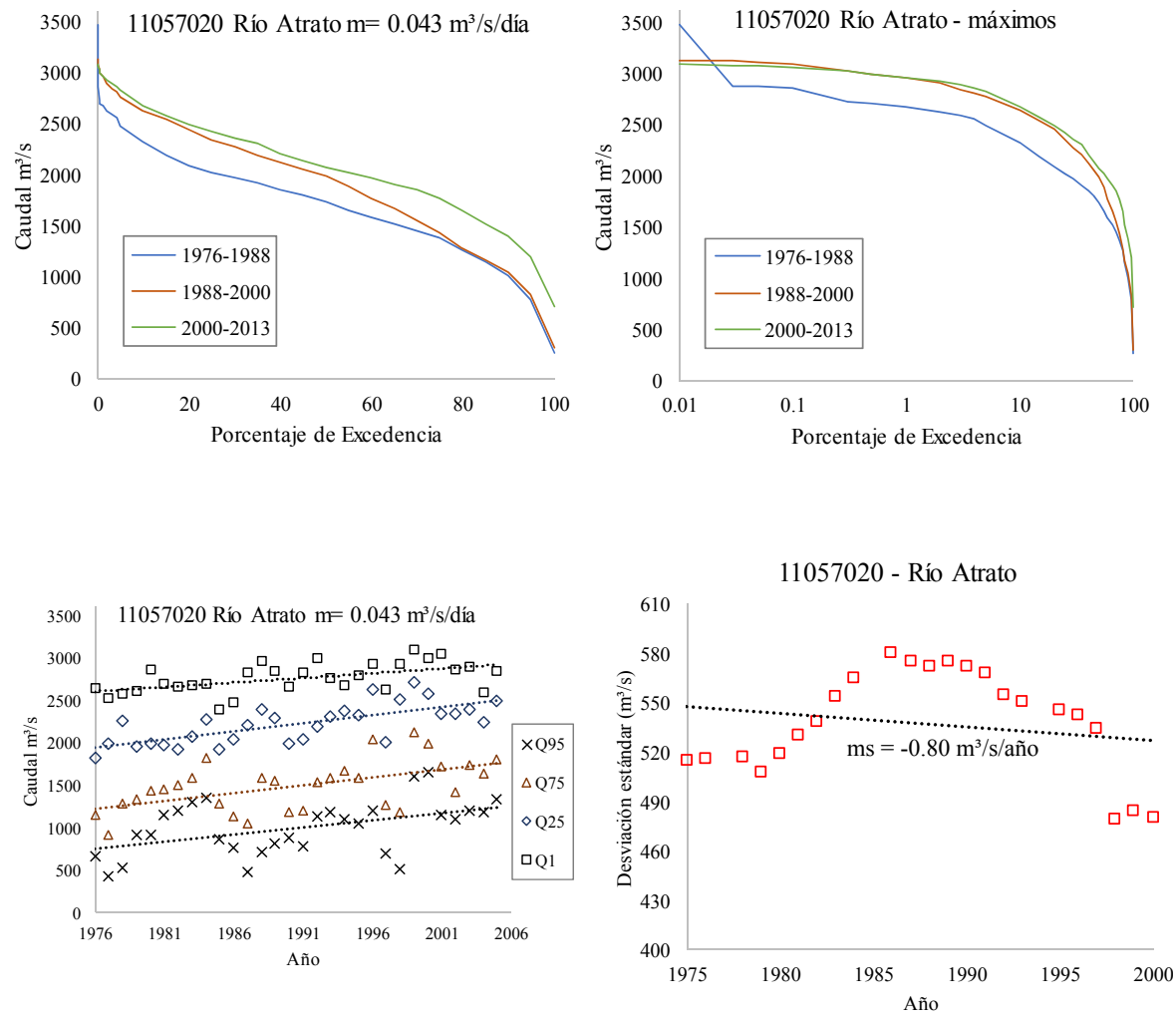


Figura 5-8. Evolución temporal de la CDC y la varianza en el río Atrato.

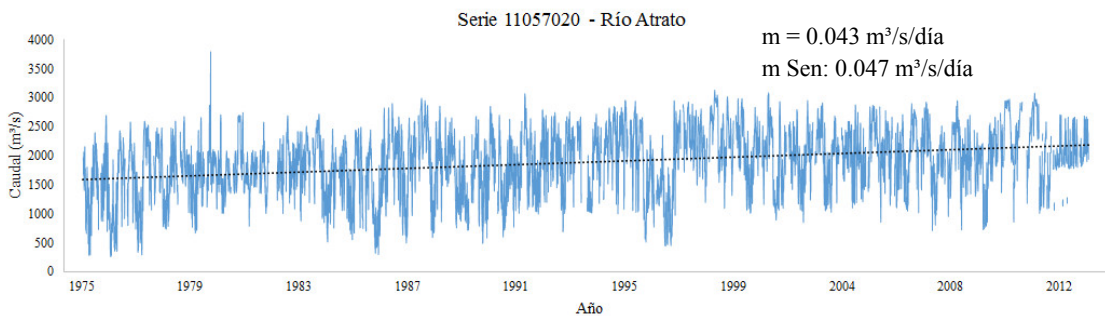


Figura 5-9. Serie temporal Río Atrato entre 1976 y 2013, ajuste lineal $m = 0.043 \text{ m}^3/\text{s}/\text{día}$.

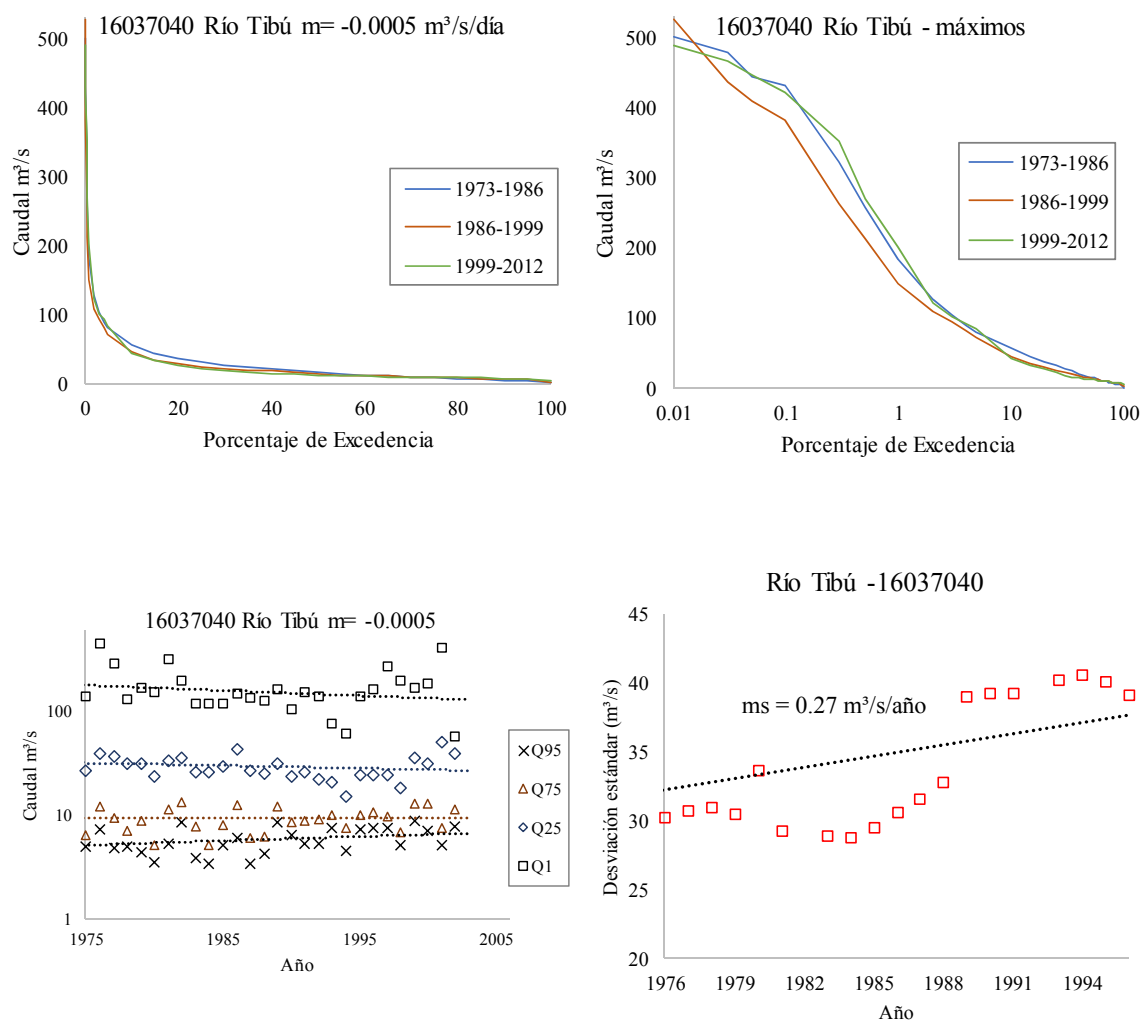
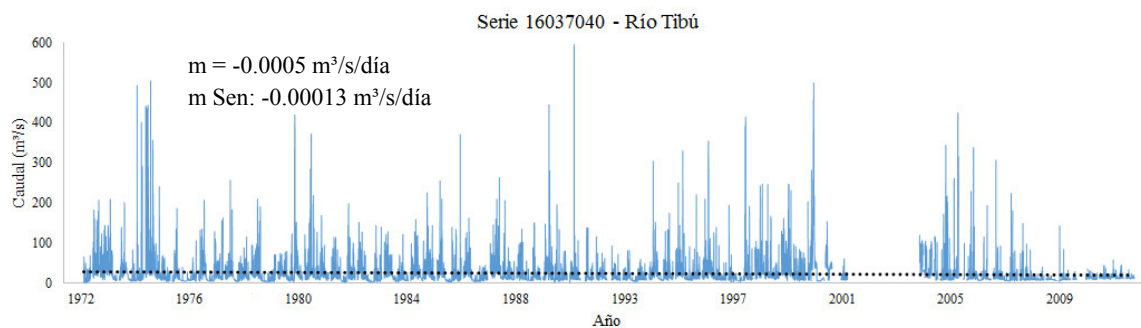


Figura 5-10. Evolución temporal de la CDC y la varianza en el río Tubú.

Figura 5-11. Serie temporal Río Tibú entre 1973 y 2012, ajuste lineal $m = -0.0005 \text{ m}^3/\text{s}/\text{día}$.

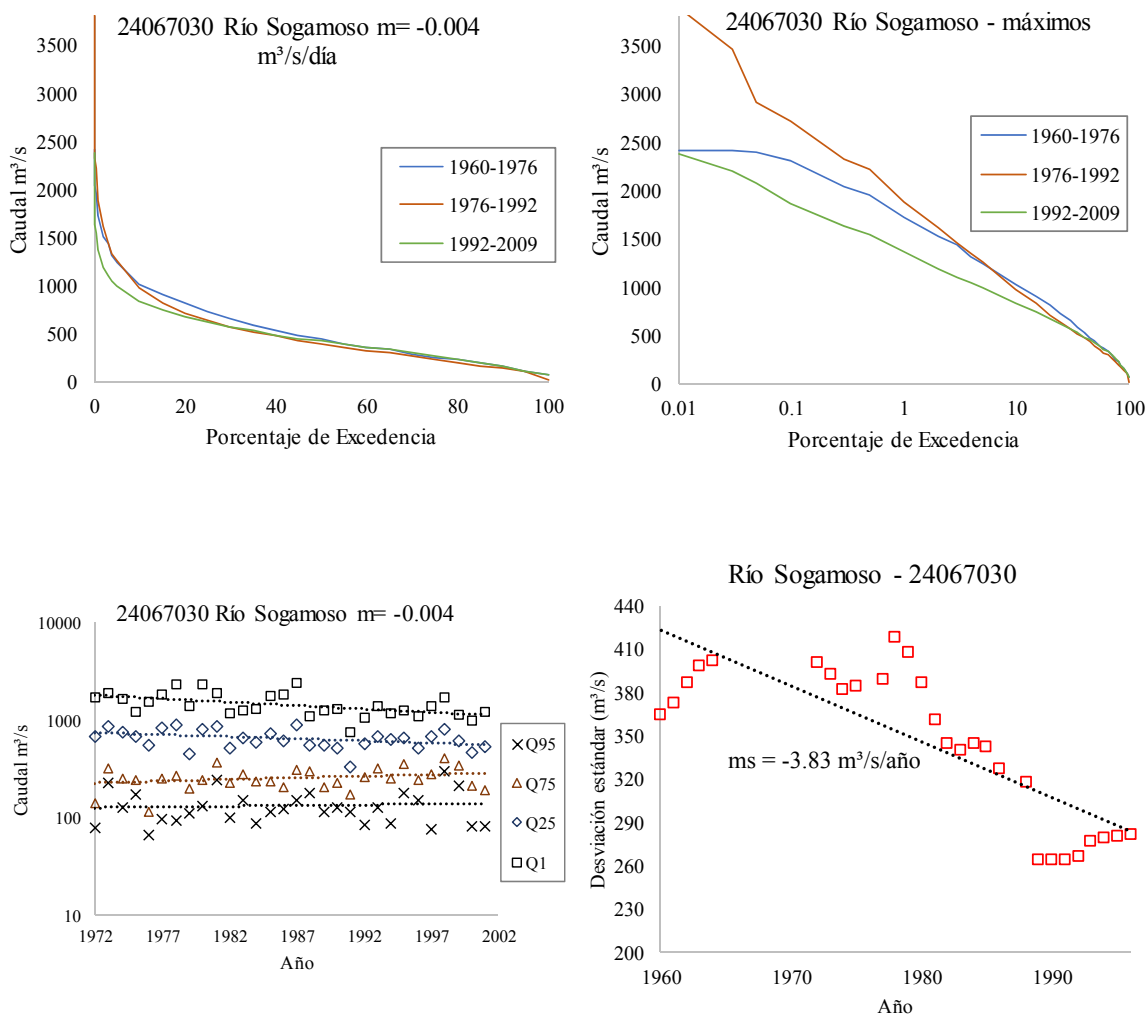


Figura 5-12. Evolución temporal de la CDC y la varianza en el río Sogamoso.

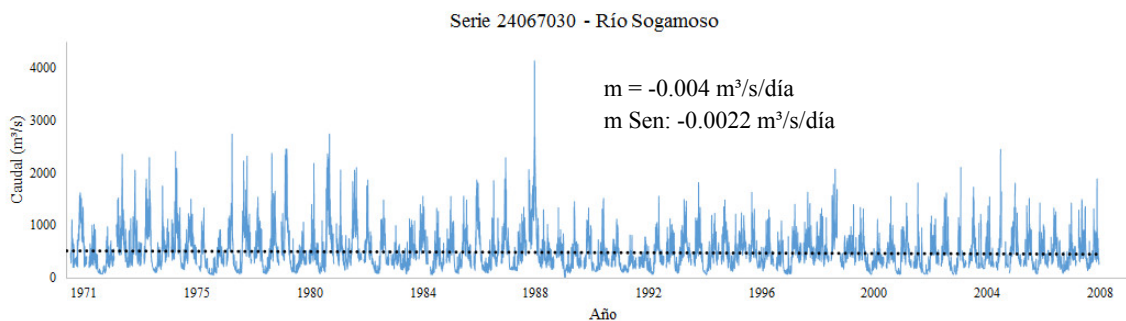


Figura 5-13. Serie temporal Río Sogamoso entre 1960 y 2009, ajuste lineal $m = -0.0048 m^3/s/día$.

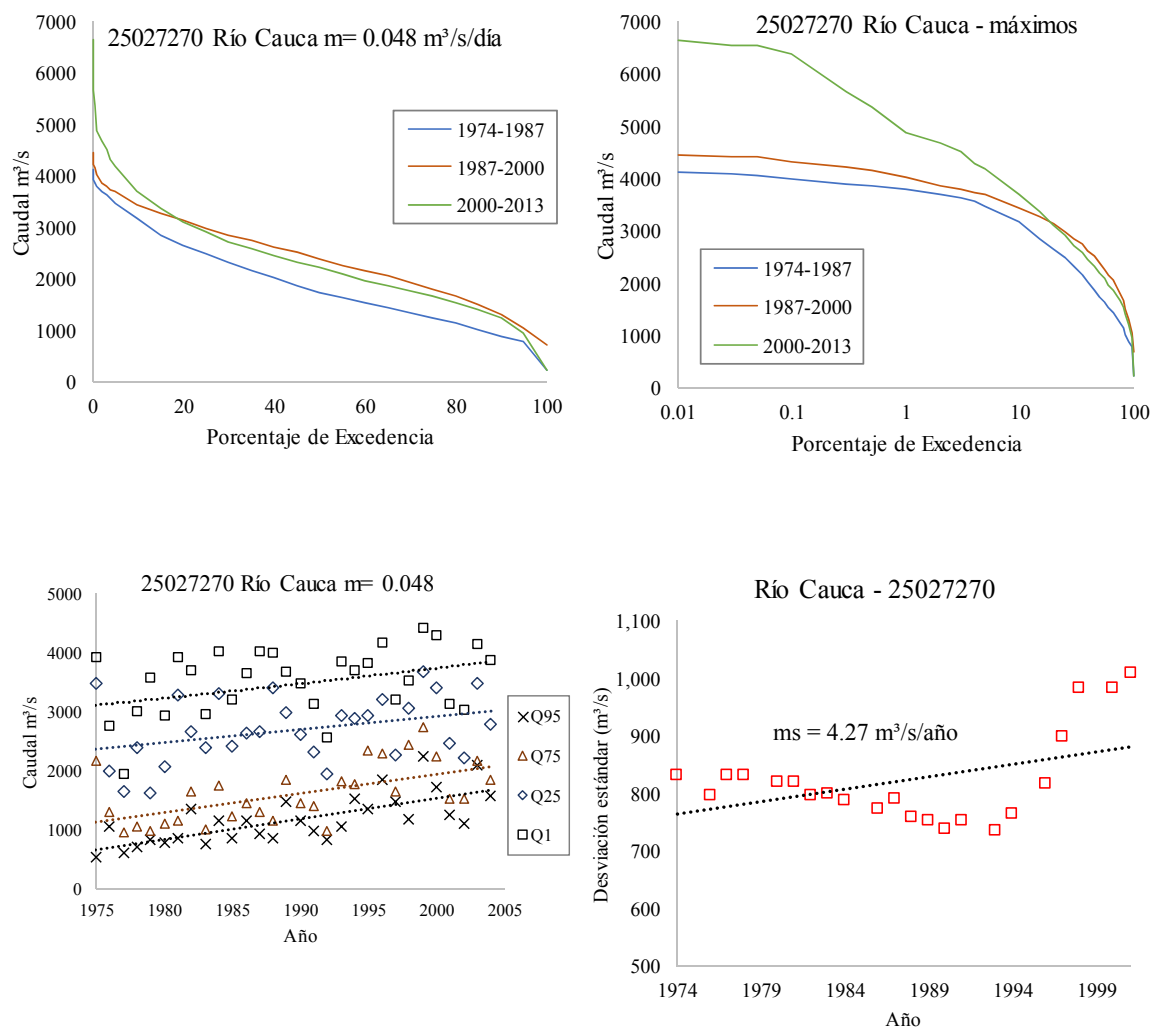
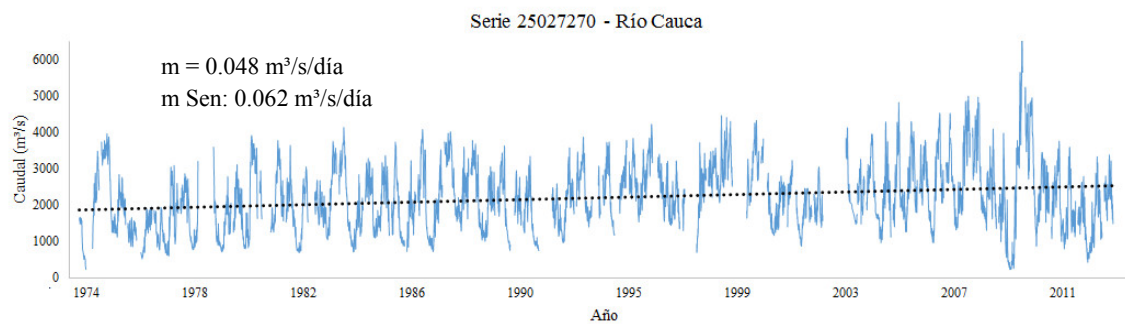


Figura 5-14. Evolución temporal de la CDC y la varianza en el río Cauca.

Figura 5-15. Serie temporal río Cauca entre 1974 y 2013, ajuste lineal $m = 0.0475 \text{ m}^3/\text{s}/\text{día}$.

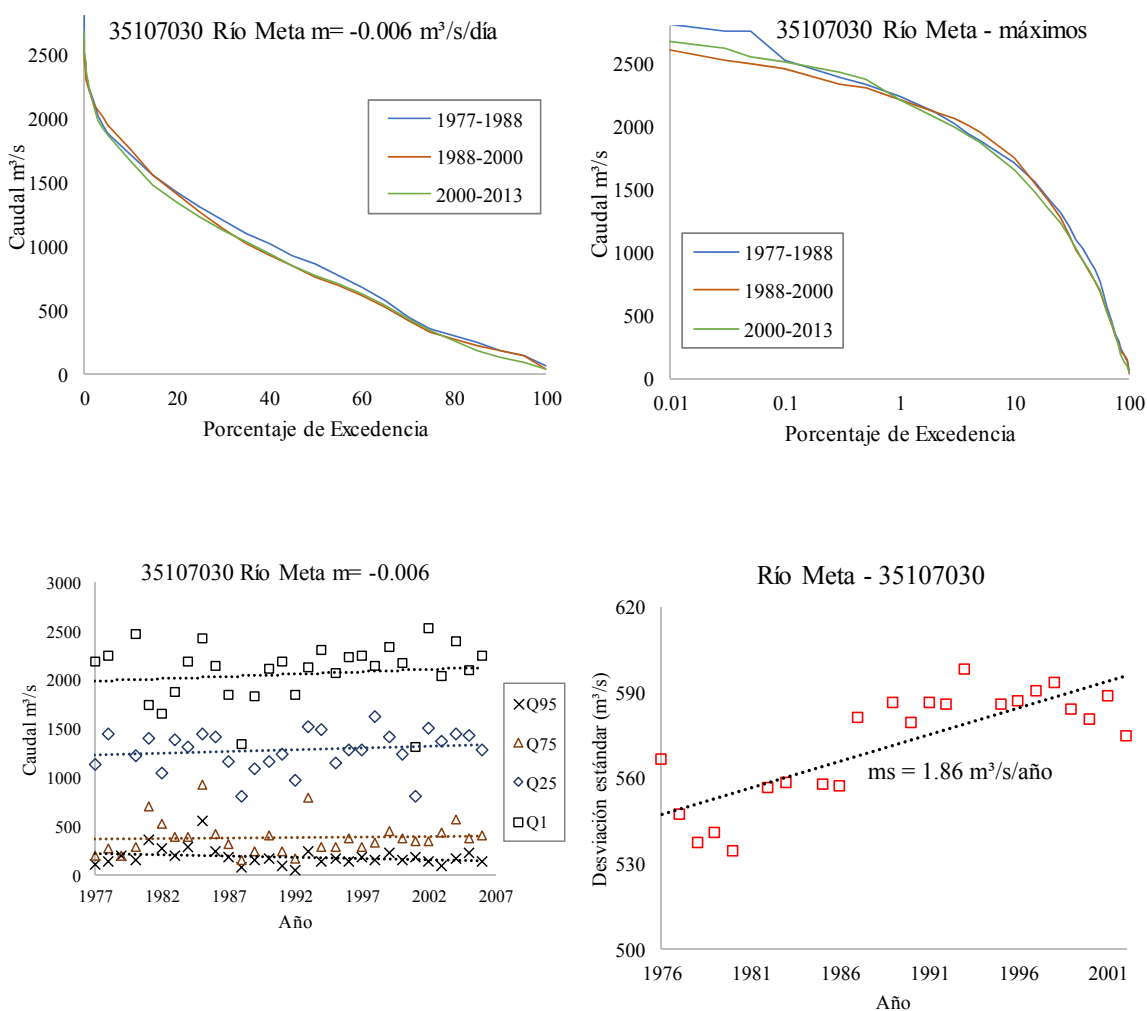


Figura 5-16. Evolución temporal de la CDC y la varianza en el río Meta.

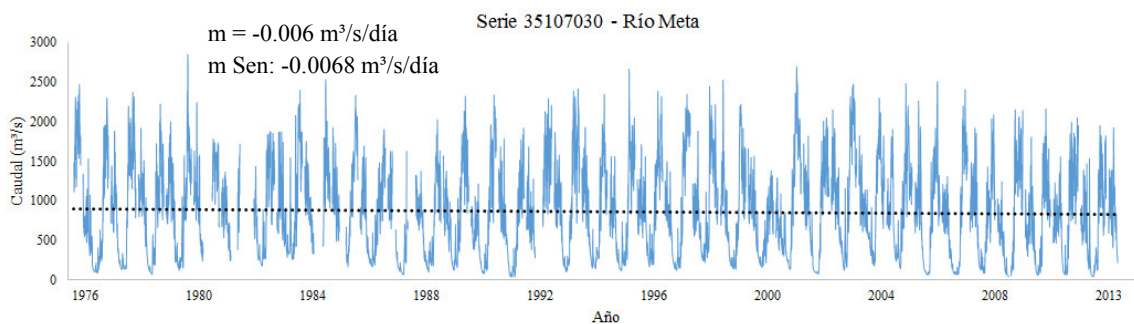


Figura 5-17. Serie temporal río Meta entre 1977 y 2013, ajuste lineal $m = -0.006 \text{ m}^3/\text{s}/\text{día}$.

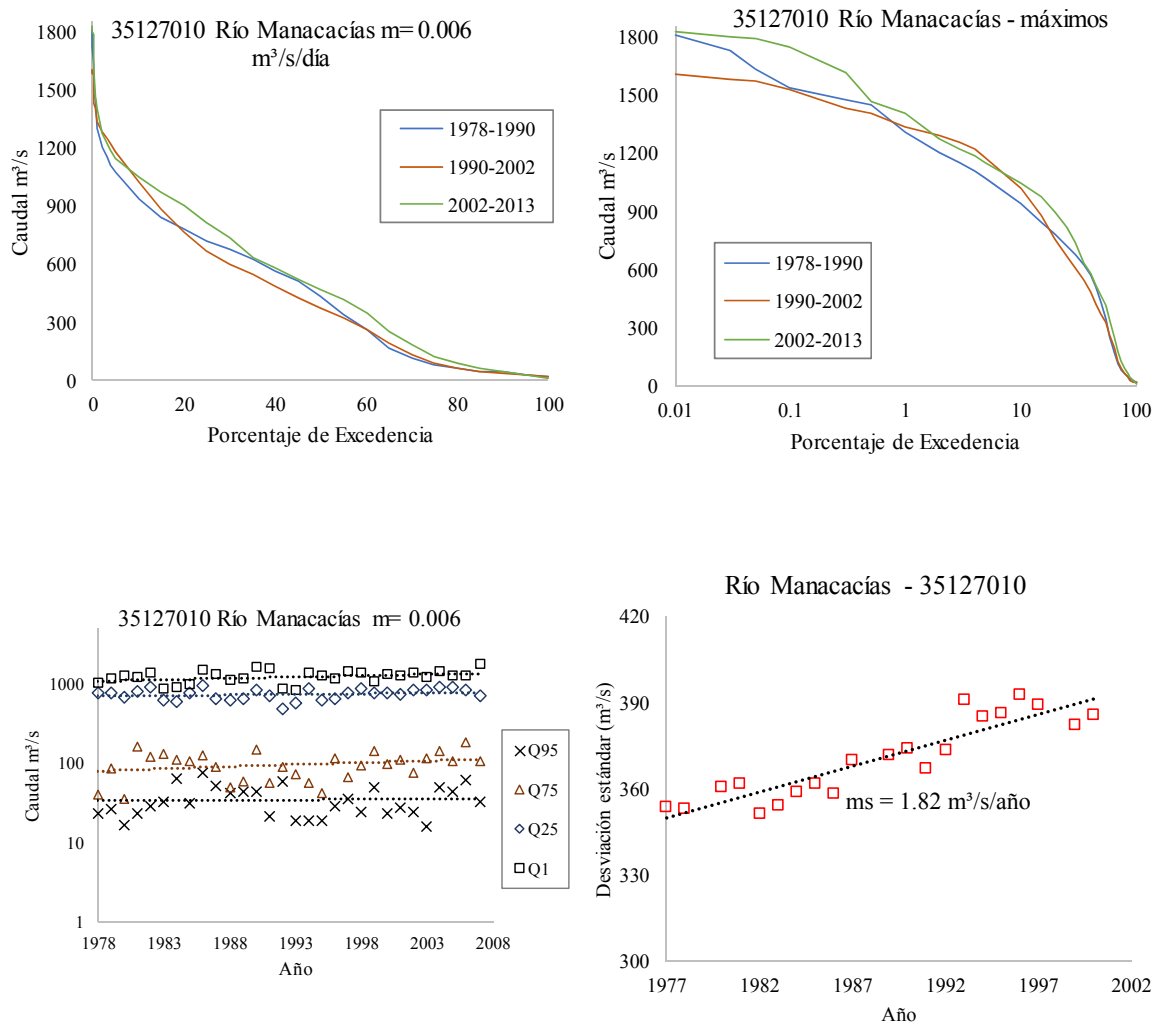
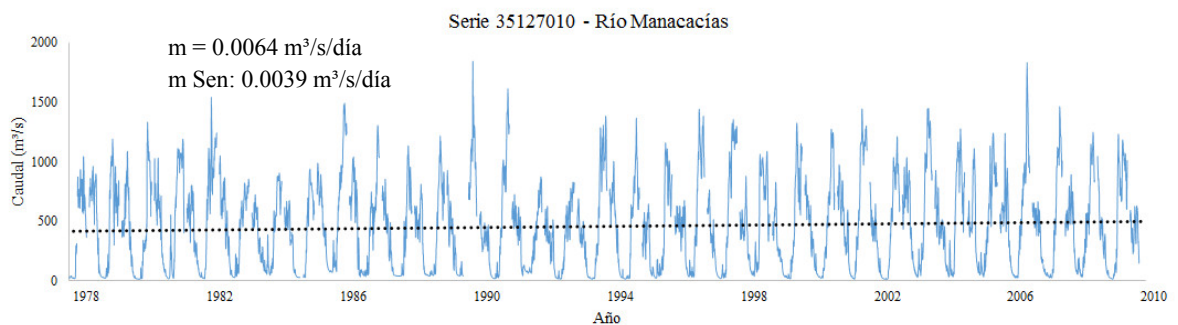


Figura 5-18. Evolución temporal de la CDC y la varianza en el río Manacacías.

Figura 5-19. Serie temporal río Manacacías entre 1978 y 2013, ajuste lineal $m=0.0064 m^3/s/día$.

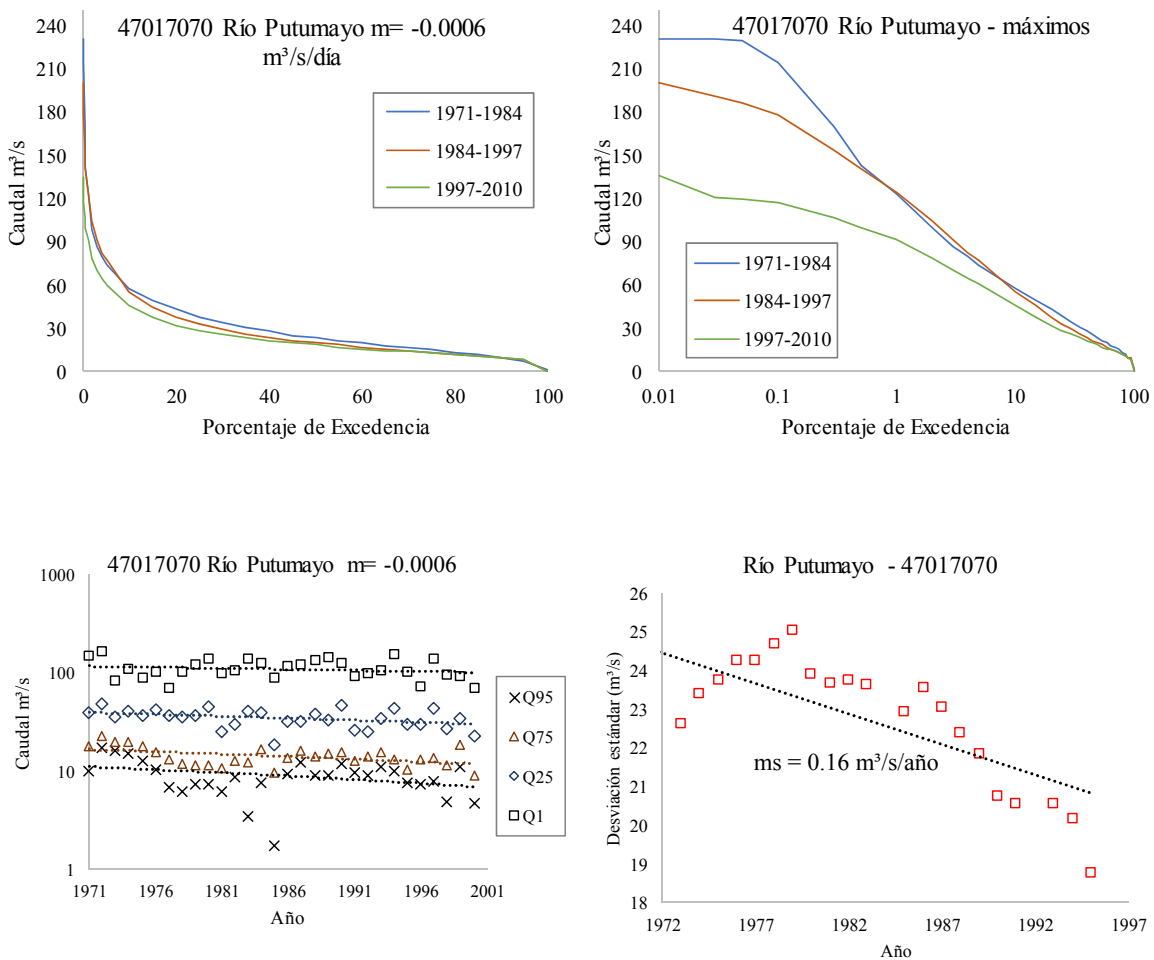


Figura 5-20. Evolución temporal de la CDC y la varianza en el río Putumayo.

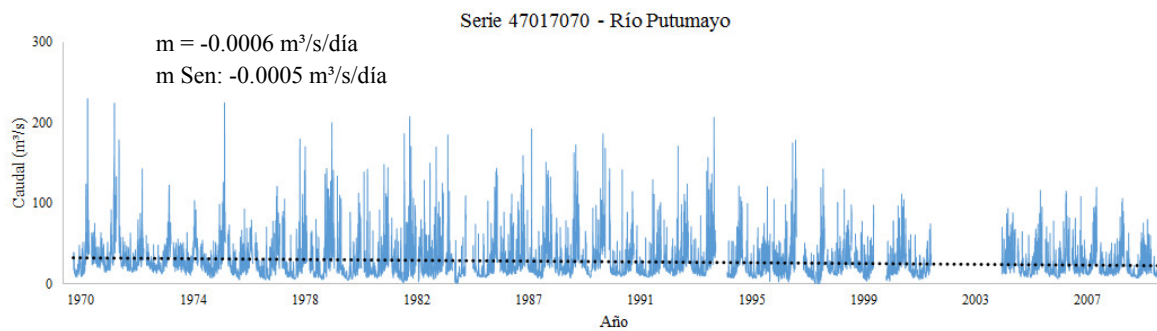


Figura 5-21. Serie temporal río Putumayo entre 1971 y 2010, ajuste lineal $m = -0.0006$ $\text{m}^3/\text{s}/\text{día}$.

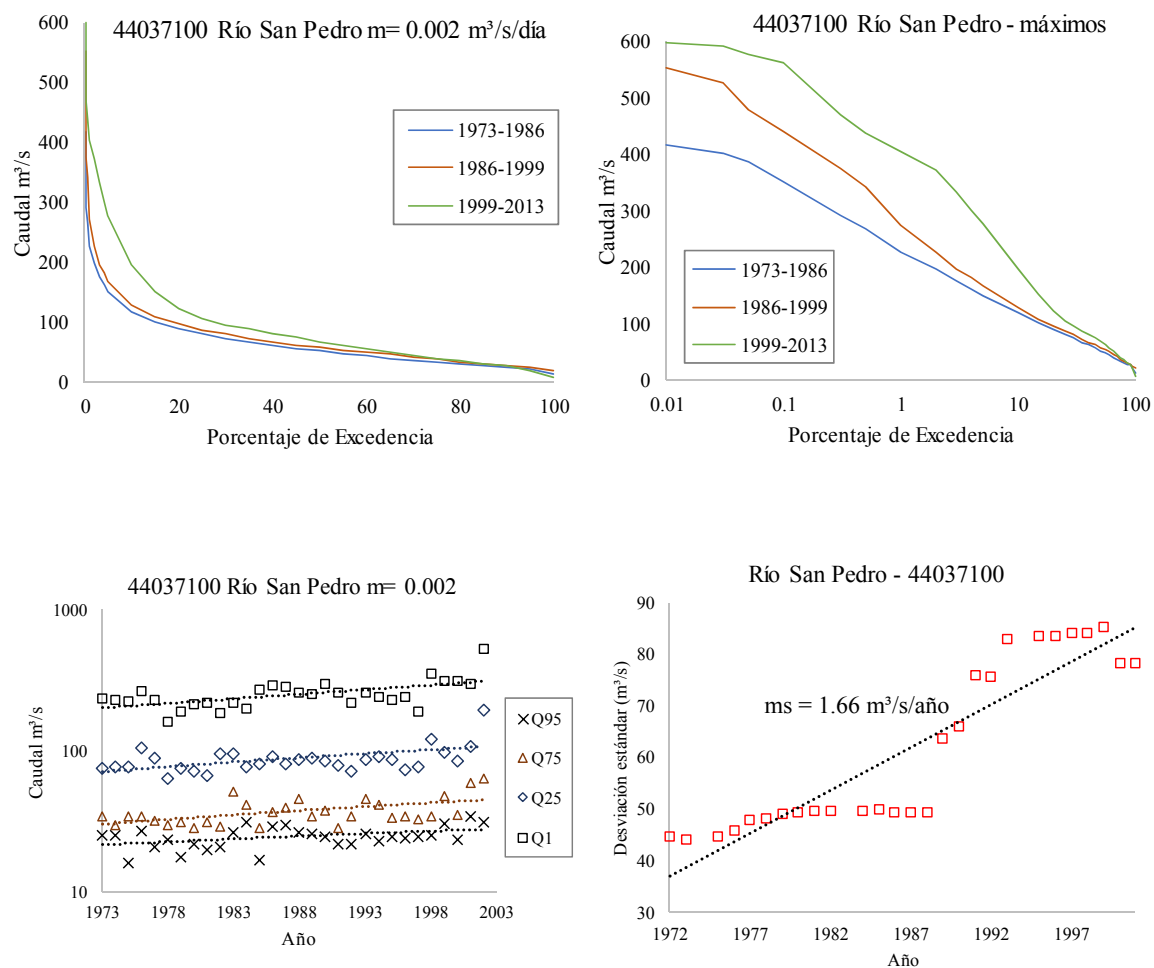


Figura 5-22. Evolución temporal de la CDC y la varianza en el río San Pedro.

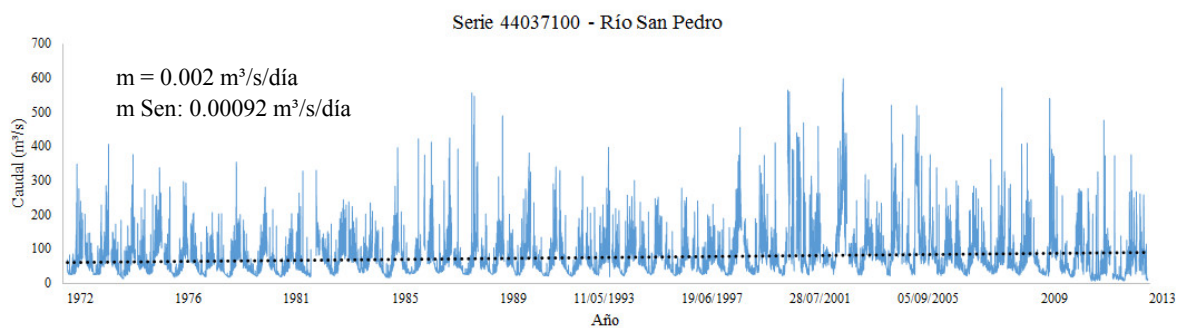


Figura 5-23. Serie temporal río San Pedro entre 1973 y 2013, ajuste lineal $m=0.002 \text{ m}^3/\text{s}/\text{día}$.

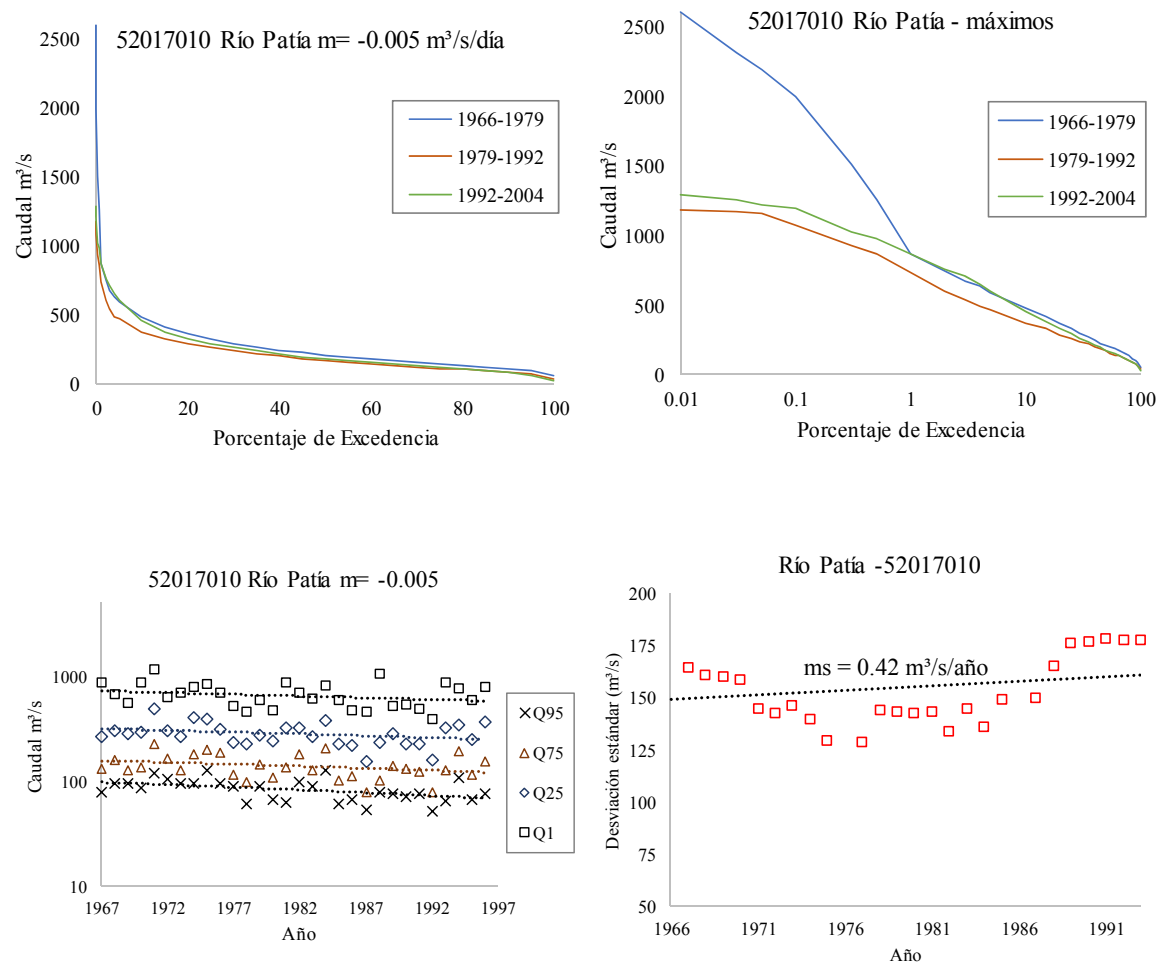


Figura 5-24. Evolución temporal de la CDC y la varianza en el río Patía.

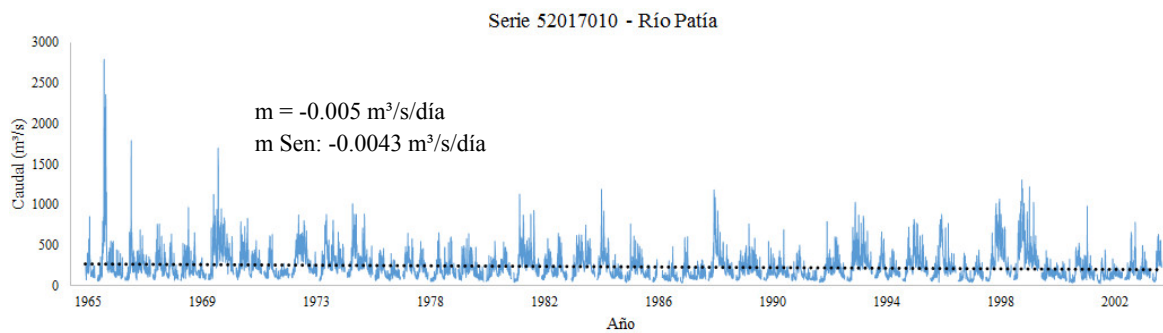


Figura 5-25. Serie temporal río Patía entre 1966 y 2004, ajuste lineal $m = -0.005 \text{ m}^3/\text{s}/\text{día}$.

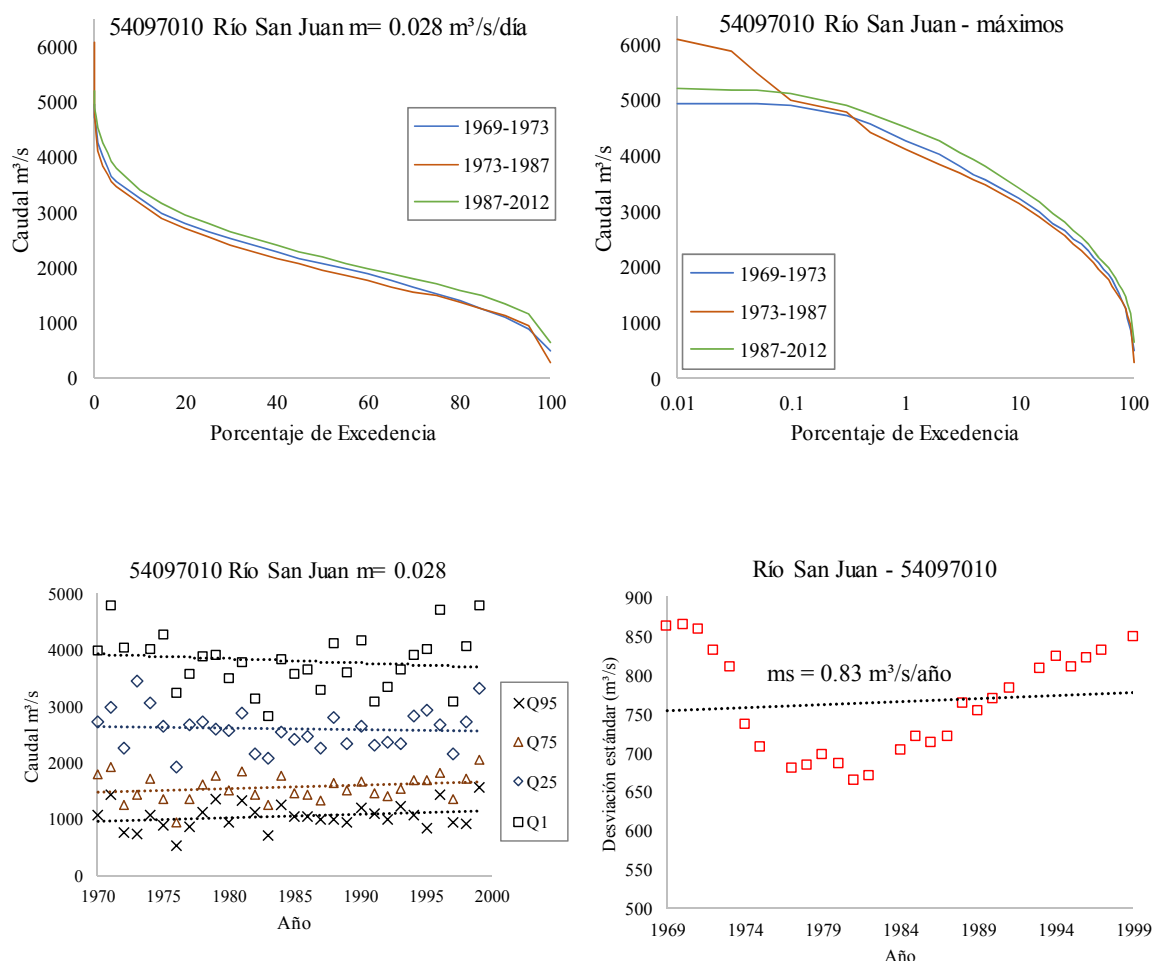
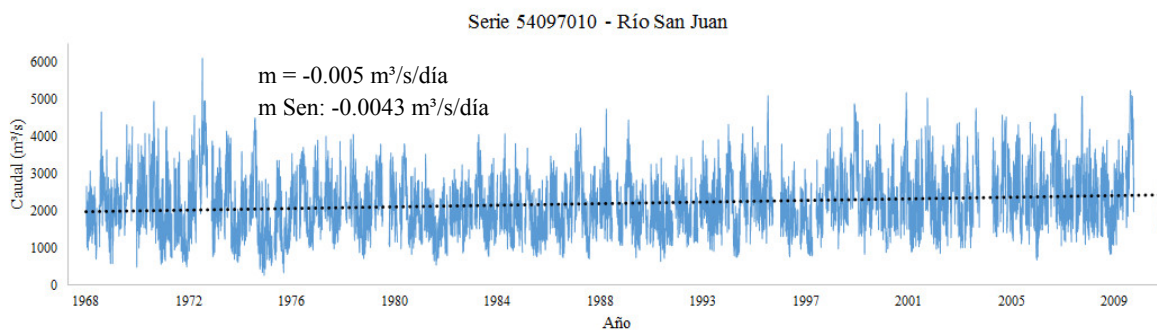


Figura 5-26 Evolución temporal de la CDC y la varianza en el río San Juan.

Figura 5-27. Serie temporal río San Juan entre 1973 y 2012, ajuste lineal $m = 0.0283 \text{ m}^3/\text{s}/\text{día}$.

La Tabla 5-1 muestra el resumen de los resultados del análisis no estacionario (valores de m, m Sen y ms) de las 10 estaciones seleccionadas.

Tabla 5-1. Resumen de resultados del análisis no estacionario.

Estación	Región	Río	Área (km ²)	Período Registro	Pend. Tend. (m ³ /s/día)	Pend. Sen (m ³ /s/día)	Pend DesvEst. (m ³ /s/año)
16037040	Caribe	Tibú	316	1973-2012	-0.0005	-0.00013	0.27
11057020	Caribe	Atrato	10566	1976-2013	0.0429	0.047	-0.80
24067030	Mag. Cauca	Sogamoso	21216	1960-2009	-0.0047	-0.0022	-3.83
25027270	Mag. Cauca	Cauca	41691	1974-2013	0.0475	0.062	4.27
35107030	Orinoquía	Meta	14401	1976-2013	-0.0058	-0.0068	1.86
35127010	Orinoquía	Manacacías	9781	1978-2013	0.0064	0.0039	1.82
47017070	Amazonía	Putumayo	474	1971-2013	-0.0006	-0.0005	0.16
44037100	Amazonía	San Pedro	636	1973-2013	0.0019	0.0009	1.66
52017010	Pacífico	Patía	8790	1966-2004	-0.0046	-0.0044	0.42
54097010	Pacífico	San Juan	12190	1969-2012	0.0283	0.032	0.83

5.4 Análisis de resultados

No fue posible realizar un análisis más detallado de cambios temporales de la CDC para las regiones Amazonía y Orinoquía en sitios aparte de las zonas de montaña, el estudio se redujo a las corrientes del pie de monte llanero y del pie de monte amazónico. Al interior tanto de la selva como de los llanos no se cuenta con series temporales seleccionadas puesto que las que se encuentran allí ubicadas (que corresponden a los ríos más grandes e importantes de las regiones) no cumplen con el mínimo de registros necesario seleccionado para el análisis, esto con excepción de algunas estaciones ubicadas en los departamentos de Meta y Caquetá. Éstas son regiones importantes para la orografía y la hidrología colombiana y representan un porcentaje importante de territorio nacional, por lo que hubiera sido de gran importancia y utilidad contar con un análisis no estacionario en estas zonas.

La máxima tendencia creciente la presentaron los ríos Atrato y Cauca ambos con más de 4% m³/s/día, esto quiere decir que sus caudales, en promedio, pasarían de 1893 y 2200 m³/s a 2050 y 2375 m³/s respectivamente en 10 años y a 2678 y 3058 m³/s respectivamente en 50 años para condiciones medias de mantenerse la tendencia. Por otra parte, el río Meta a los 14400 km² de área de drenaje con una pendiente de tendencia de -0.0058 m³/s/día y un

caudal medio de 861 m³/s, se proyecta un caudal medio de 839.8 m³/s en un plazo de 10 años y un caudal medio de 755.2 m³/s en un plazo de 50 años, siendo esta la mayor tendencia negativa se nota que el cambio no es tan brusco, las tendencias crecientes son considerablemente más significativas (un orden de magnitud más). Es posible que estos valores de tendencias se deban a cambios de tipo morfométrico como sedimentación (agradación) en la sección de aforo, más no necesariamente a una evolución en las condiciones climáticas y fisiológicas de la cuenca.

Analizando tendencias anuales de las diferentes magnitudes de caudales, en los casos de los ríos Cauca, Atrato y Manacacías, las tendencias son crecientes para todos los percentiles característicos analizados, es decir, desde los caudales mínimos hasta los caudales máximos, todos presentan un comportamiento creciente. En el caso de los ríos Patía y Tibú, en Nariño y Norte de Santander respectivamente, se observa una tendencia decreciente generalizada, pero principalmente ocasionada por un decrecimiento en los valores pico de inundaciones. Los caudales del río Sogamoso, en Santander, presentan tendencias decrecientes con excepción del percentil 75 (caudales medios - mínimos) el cual presenta un comportamiento temporal creciente. El río Meta presenta una pendiente negativa de forma general, sin embargo los valores de máximos y medios – máximos presentan tendencias crecientes. La tendencia creciente generalizada que presentan los caudales del río San Pedro en el Caquetá, se dan principalmente por un incremento gradual en el valor de los caudales máximos, mientras que el río San Juan, entre los departamentos de Cauca y Chocó, a pesar de la tendencias decrecientes en los valores máximos, la pendiente de la regresión lineal de la serie, debido al comportamiento de los caudales medios, es positiva.

Los cambios temporales en la parte de máximos de la CDC pueden estar asociados algunas veces a un único pico de creciente, como se puede ver por ejemplo en la Figura 5-24 y en la Figura 5-25 con el río Patía, en este caso los eventos de creciente ocurridos en 1966 y en 1967 provocan un cambio significativo en los caudales $Q_{0.3}$, $Q_{0.1}$, $Q_{0.03}$ y $Q_{0.01}$ respecto a los otros períodos analizados. En el mismo caso del río Patía, se observa que los caudales medios y mínimos son mayores entre 1992 y 2004 que entre 1979 y 1992, aun así, los picos de caudales ocurridos entre 1966 y 1979 hacen que la tendencia sea negativa y la serie decreciente. En el río San Juan en cambio, la tendencia creciente es de forma homogénea, esto quiere decir que tanto en máximos, medios y mínimos los caudales van aumentando en cada período (Figura 5-26 y Figura 5-27).

La pendiente de Sen y la pendiente de la recta de ajuste presentan una relación importante R^2 de 0.87. Ambos valores provienen de diferentes metodologías de cálculo, la pendiente de Sen está dada por la mediana de las pendientes dato a dato de la serie, mientras que la pendiente de ajuste se obtiene a partir de la recta obtenida por mínimos cuadrados, sin embargo, puede concluirse que cualquiera de los dos criterios es igual de válido para realizar estimación de tendencias lineales a largo plazo.

Se tienen casos en los cuales, el valor de la desviación cambia muy poco en comparación con la magnitud del caudal medio, por ejemplo los ríos Atrato ($ms = -0.8 \text{ m}^3/s/año$), Putumayo ($ms = 0.16 \text{ m}^3/s/año$) y San Juan ($ms = -0.8 \text{ m}^3/s/año$). Se tienen casos también en los cuales los valores de la desviación oscilan a través del período de estudio, como lo son por ejemplo, los ríos Atrato, Cauca, Putumayo, Patía y San Juan. Estas oscilaciones, hacen que la forma de tendencia lineal, no sea la más adecuada para su representación.

Al observar tendencias en la varianza, para las 10 series temporales seleccionadas, se tiene que no hay una relación clara entre el valor de pendiente del ajuste a la desviación estándar y la pendiente del ajuste de los caudales. Por ejemplo, en el río Atrato, mientras el análisis de datos indica que los caudales en promedio incrementan en aproximadamente $15 \text{ m}^3/s$ en un año, la desviación estándar disminuye $0.8 \text{ m}^3/s$ en ese mismo período, algo similar ocurre en los ríos Tibú, Meta y Patía. Los cambios en la varianza también se dan en su mayoría de forma creciente, con excepción de los ríos Atrato y Putumayo, cuya tendencia no es significativa (decrecen con una tasa menor a $1 \text{ m}^3/s/año$) y el río Sogamoso, única corriente con una tendencia decreciente considerable ($-3.83 \text{ m}^3/s/año$).

6 Conclusiones y Recomendaciones

6.1 Conclusiones

A pesar del enorme déficit de información hidrológica en Colombia, tanto espacial como temporal, fue posible extrapolar en este trabajo condiciones para la estimación del régimen de caudales en los sitios sin información. Se asume que las estaciones de medición que fueron utilizadas y sus respectivas cuencas hidrográficas abarcan un espectro lo suficientemente amplio en características como: tamaño, forma, caudales medios, condiciones climáticas, tipos de suelo entre otras. Esto hizo posible que las aproximaciones pudieran ser realizadas con probabilidades de éxito similares para ríos en todo el país de diferentes tamaños y con diferentes condiciones.

Se produjo una serie de ecuaciones funcionales que permiten la estimación de la curva de duración de caudales diaria para cuencas con información escasa en Colombia. Los atributos utilizados en este trabajo se encuentran disponibles para descargar de forma gratuita, se necesita entonces seleccionar la cuenca a la cual se le realizará la aproximación, definir si se estimará mediante ecuaciones de clúster o ecuaciones de sub regiones, ubicarla en un clúster o en una sub región, definir los atributos necesarios, descargar la información y extraerlos mediante sistemas de información geográfica. Las unidades de los atributos se encuentran especificadas. En caso de que se escojan las ecuaciones de clúster, éste puede ubicarse espacialmente con el mapa de regiones hidrológicas o calculando las distancias euclidianas a los centroides geométricos de cada clúster (Anexo B), para realizar esto último, es necesario contar con los 25 atributos.

La variabilidad espacial de las diferentes variables hidrológicas del país se evidencia a lo largo de este trabajo, de manera tradicional se han definido cinco sub regiones predominantes, sin embargo, son muchas las variables relacionadas con el comportamiento de una cuenca y las mismas se distribuyen de forma heterogénea en el territorio, lo que ocasiona que el problema sea más complejo. El criterio de vecindad o cercanía, no tuvo

ninguna relevancia a la hora de delimitar regiones hidrológicas en este trabajo, por lo que puede concluirse que, dos cuencas, no por el hecho de ser vecinas, estarán necesariamente ubicadas dentro de una región hidrológica y por lo tanto sus respuestas no tienen por qué ser similares.

La regresión lineal múltiple, tal y como fue planteada en este trabajo, representa una nueva manera de estimar el régimen de caudales en los ríos de Colombia. Según los resultados obtenidos en las diferentes validaciones, se tienen aproximaciones con errores porcentuales bajos para el régimen de caudales medios, generalmente, de menos del 10%, los valores extremos tanto para mínimos como para máximos continúan presentando incertidumbres mayores, hasta 50% de error relativo promedio en los percentiles correspondientes a los mínimos (especialmente en el Q_{100}) y aproximadamente 30% de error relativo medio en los percentiles correspondientes a caudales máximos. Los altos errores obtenidos en las aproximaciones del Q_{100} son atribuidos a las magnitudes mismas de este caudal característico, por tratarse de un error relativo, al hacerse la relación entre números cercanos a cero, se obtienen valores más altos de error relativo porcentual.

El tamaño de una cuenca es muy influyente a la hora de realizar la estimación del régimen de caudales aunque éste sea adimensional, pues la forma misma de la curva de duración está fuertemente relacionada con éste. Los modelos aquí planteados, tanto para la agrupación y la ubicación de cuencas dentro de un determinado clúster, como para la estimación del régimen de caudales, tienen en cuenta la extensión geográfica de las cuencas, esto debido a que el parámetro área de drenaje no es el único relacionado con el tamaño, las variables perímetro de la cuenca, índice de compacidad, densidad de falla, longitud del cauce principal, y desnivel presentan estrechas relaciones con el área (con valores de R mayores o cercanos a 50%) por lo que se intuye, que tienen en cuenta de manera explícita el tamaño. En el Anexo D se muestran las dispersiones de cada uno de estos atributos vs el área de drenaje para las 655 cuencas.

En el análisis de correlaciones en las regresiones lineales múltiples, se observa un importante valle en los valores de R^2 correspondientes a la estimación del Q_{20} y los percentiles cercanos (Figura 4-1 y Figura 4-3). A pesar de que en los mínimos también se observan valores bajos de R^2 , (Q_{90} y Q_{100}) estos pueden explicarse mediante la incertidumbre asociada a las mismas magnitudes, pues coinciden las bajas correlaciones con los altos errores relativos en la validación (Tabla 4-11 y Tabla 4-12), el valle ubicado en el percentil 20 es más complejo de interpretar pues éste corresponde aún al régimen medio de descargas. Sin embargo, el

ejemplo mostrado en la Figura 3-10 señala que este percentil, se ubica en una transición entre valores negativos y positivos de R , por lo que éstos son próximos a cero en esta zona.

Los resultados de la validación varían conforme cual sea la estación utilizada para este fin, se pueden obtener muy buenas aproximaciones en un clúster que haya obtenido la calificación de regular, así como altos errores porcentuales en clúster catalogados como muy buenos, esto puede deberse a que los atributos con los que se estiman algunos percentiles no se encuentran dentro del rango de valores de los mismos con los que fueron calibradas las ecuaciones, esto puede generar grandes diferencias entre valor observado y estimado.

El modelo planteado, estima el régimen de caudales adimensionales Q^* , el cual debe ser multiplicado por el caudal medio a largo plazo para conocer la curva de duración de caudales en m^3/s . Para todo el territorio nacional también existe la posibilidad de estimar las variables precipitación y evapotranspiración real con las cuales se puede calcular dicho caudal, sin embargo, esto es otra fuente de error, pues el balance hídrico es estimado a gran escala con las variables precipitación y evapotranspiración real, esta última, con muy pocos registros en el país y con altas incertidumbres asociadas en su estimación.

De forma generalizada, los cambios en el régimen de caudales en Colombia se dan principalmente con tendencias positivas (caudales crecientes), según esto, los picos de inundaciones serían cada vez más frecuentes y las magnitudes de los caudales máximos serían mayores, esto sería coherente con lo reportado en la literatura de análisis de no estacionario de variables hidrológicas desarrollada en el numeral 1.3.3. Además de obtener tendencias crecientes en la mayoría de las estaciones, en las corrientes analizadas también se encontró que por lo general la desviación estándar tiende a incrementar, lo que indica, que las cuencas tienden a regular menos sus caudales, cada vez los valores tanto mínimos como máximos se encuentran más lejos del valor promedio.

6.2 Recomendaciones y trabajo futuro

Este trabajo puede realizarse con una mejor resolución espacial del modelo de elevación digital *DEM*, hay disponibles planchas con información topográfica digital gratuita en resoluciones de 30 (*SRTM*) y 12.5 (*Alos Palsar*) metros, si bien incrementa de forma considerable el gasto computacional, se podría realizar el análisis con mayor detalle, sobre todo para las cuencas de menor tamaño (menores a 100 km) y sus respectivos atributos dependientes de la topografía y el trazado.

La información cartográfica utilizada también fue a escala ‘gruesa’ (1:100.000) pues es la que se encuentra disponible de forma gratuita para todo el país, sin embargo existen mapas geológicos, geomorfológicos, de cobertura y uso del suelo con más detalle para algunas regiones del país, con lo que se podría asignar los atributos a las cuencas correspondientes a esta información con un mayor nivel de detalle.

El mapa de evapotranspiración potencial *ETP* utilizado en este trabajo, fue calculado a partir de la ecuación de Cenicafé, ésta depende únicamente de la elevación por lo que se hace sencilla su estimación, sin embargo, existen otras metodologías para la obtención tanto de *ETP* como de *ETR* (Evapotranspiración Real), además de datos en tierra con mediciones de tanques o lisímetros con los cuales se puede realizar una interpolación espacial e incluir todas estas variables a la lista de atributos.

Además del *K-means*, existen otras metodologías para agrupar cuencas, entre ellas se encuentran un clúster jerárquico, las curvas de Andrews, dendogramas, entre otros, con el fin de comparar los resultados de regionalización obtenidos a partir de clústers obtenidos mediante diferentes metodologías de agrupación de cuencas en Colombia.

Para la agrupación de cuencas sin calibrar en Colombia, se puede utilizar un trazado de cuencas con un umbral de área menor, en este trabajo se agruparon cerca de 400 unidades hidrológicas de todo el país, sin embargo es posible realizar la misma agrupación pero a un número mucho más grande de sub cuencas. En este aspecto, se espera que los resultados sean diferentes, pues muchos de los atributos están relacionados con el tamaño de las cuencas, lo ideal sería, que la distribución de tamaños a regionalizar, sea igual para las cuencas calibradas que para las cuencas sin calibrar.

En este trabajo no fue posible realizar diferentes agrupaciones de cuencas o clústers para las diferentes magnitudes de caudales o tramos de la CDC pues las variables que mejor se correlacionaban con los diferentes caudales característicos seleccionados eran casi las mismas, esto a pesar de que los valores mismos de los coeficientes de correlación varían a lo largo del porcentaje de excedencia (Figura 3-10). Se recomienda a futuro, una revisión más detallada para este análisis.

La realización de proyecciones para diferentes escenarios tendenciales a la regionalización de las CDC con base en las tendencias de cambio a largo plazo obtenidas en el análisis no estacionario de la curva de duración, pueden hacerse, las regresiones lineales múltiples pero calibradas con los percentiles de las CDC reales proyectados a, por ejemplo, cien años.

Hace falta un estudio a profundidad de cuáles son las razones o los motivos por los cuales cambia a largo plazo el régimen de caudales de algunos ríos de Colombia, si bien se conoce que el cambio climático es un factor determinante en este ámbito, también las variables asociadas a coberturas y usos del suelo explican en parte el comportamiento de los caudales, por lo tanto, los cambios antrópicos a los cuales éstas son sometidas también deben ser tenidos en cuenta a la hora de estudiar las alteraciones temporales en una CDC.

A. Resumen de Resultados Matriz de Correlaciones

Valores de R para atributos vs caudales adimensionales

Atributo	Coeficiente de Correlación		
	Máximo	Medio	Mínimo
ADre	0.300	0.043	-0.291
Perim	0.392	0.053	-0.370
Komp	0.319	0.037	-0.299
%Agr	0.113	-0.014	-0.106
%Bosq	0.344	0.042	-0.324
%Urb	0.096	0.006	-0.115
DenF	0.128	0.018	-0.106
DenD	0.173	0.016	-0.174
%lim	0.168	-0.010	-0.054
%Aren	0.220	-0.027	-0.116
%Arc	0.220	0.017	-0.166
ETPa	0.123	-0.018	-0.119
Emax	0.399	0.050	-0.351
Emed	0.223	0.016	-0.122
Emin	0.085	-0.012	-0.119
Desn	0.402	0.056	-0.386
LonC	0.387	0.052	-0.367
Pend	0.305	0.030	-0.178
Pmáx	0.189	0.022	-0.178
Pmed	0.310	0.038	-0.290
Tmed	0.120	-0.015	-0.109
H10	0.141	0.020	-0.099
H25	0.182	0.019	-0.103
H50	0.179	0.018	-0.099
H75	0.170	0.017	-0.091

Q_p	Coeficiente de Correlación		
	Máximo	Medio	Mínimo
100	0.294	0.121	-0.123
95	0.389	0.127	-0.230
90	0.397	0.130	-0.239
85	0.396	0.131	-0.235
80	0.399	0.135	-0.230
75	0.398	0.138	-0.212
70	0.391	0.143	-0.188
65	0.401	0.151	-0.164
60	0.402	0.157	-0.138
55	0.398	0.160	-0.119
50	0.391	0.160	-0.104
45	0.392	0.159	-0.106
40	0.384	0.154	-0.104
35	0.368	0.145	-0.107
30	0.338	0.128	-0.108
25	0.285	0.103	-0.133
20	0.168	0.040	-0.146
15	0.217	-0.038	-0.172
10	0.220	-0.112	-0.273
5	0.145	-0.152	-0.370
4	0.121	-0.149	-0.368
3	0.104	-0.150	-0.374
2	0.103	-0.152	-0.386
1	0.111	-0.149	-0.382
0.5	0.113	-0.143	-0.365
0.3	0.107	-0.131	-0.345
0.1	0.119	-0.117	-0.314
0.05	0.123	-0.111	-0.311
0.03	0.122	-0.110	-0.308
0.01	0.107	-0.088	-0.247

Valores de R para atributos vs caudales normalizados

Atributo	Coeficiente de Correlación		
	Máximo	Medio	Mínimo
ADre	-0.004	-0.041	-0.098
Perim	-0.008	-0.060	-0.148
Komp	-0.028	-0.062	-0.162
%Agr	-0.036	-0.049	-0.061
%Bosq	0.098	0.060	0.030
%Urb	-0.033	-0.039	-0.043
DenF	-0.054	-0.070	-0.084
DenD	-0.005	-0.052	-0.108
%lim	-0.007	-0.013	-0.028
%Aren	0.048	0.035	-0.003
%Arc	0.012	-0.009	-0.048
ETPa	0.146	0.127	0.122
E _{max}	-0.099	-0.182	-0.248
E _{med}	-0.068	-0.123	-0.140
E _{min}	-0.029	-0.057	-0.075
Desn	-0.079	-0.142	-0.220
LonC	-0.010	-0.060	-0.145
Pend	-0.002	-0.046	-0.083
P _{máx}	0.111	0.092	0.079
P _{med}	0.134	0.104	0.089
T _{med}	0.137	0.122	0.121
H10	-0.007	-0.013	-0.016
H25	-0.009	-0.017	-0.020
H50	-0.002	-0.014	-0.017
H75	0.018	0.005	-0.004

Q _p	Coeficiente de Correlación		
	Máximo	Medio	Mínimo
100	0.072	-0.009	-0.099
95	0.123	-0.009	-0.113
90	0.117	-0.010	-0.120
85	0.125	-0.010	-0.131
80	0.134	-0.011	-0.139
75	0.131	-0.012	-0.150
70	0.127	-0.013	-0.154
65	0.130	-0.013	-0.158
60	0.134	-0.014	-0.163

Q_p	Coeficiente de Correlación		
	Máximo	Medio	Mínimo
55	0.134	-0.014	-0.165
50	0.137	-0.015	-0.169
45	0.141	-0.016	-0.175
40	0.143	-0.016	-0.178
35	0.144	-0.017	-0.181
30	0.145	-0.018	-0.185
25	0.146	-0.019	-0.189
20	0.138	-0.021	-0.188
15	0.139	-0.022	-0.191
10	0.139	-0.025	-0.196
5	0.136	-0.028	-0.201
4	0.134	-0.029	-0.202
3	0.133	-0.030	-0.204
2	0.130	-0.032	-0.207
1	0.127	-0.034	-0.210
0.5	0.122	-0.037	-0.212
0.3	0.123	-0.039	-0.217
0.1	0.123	-0.043	-0.225
0.05	0.126	-0.047	-0.237
0.03	0.125	-0.048	-0.239
0.01	0.130	-0.052	-0.248

Valores promedio de R en máximos y mínimos entre atributos en cada clúster

Clúster	R	
	Promedio máximos	Promedio mínimos
1	0.326	-0.319
2	0.699	-0.686
3	0.533	-0.489
4	0.249	-0.230
5	0.337	-0.331
6	0.304	-0.311
7	0.382	-0.381
8	0.260	-0.263
9	0.485	-0.486
10	0.326	-0.330

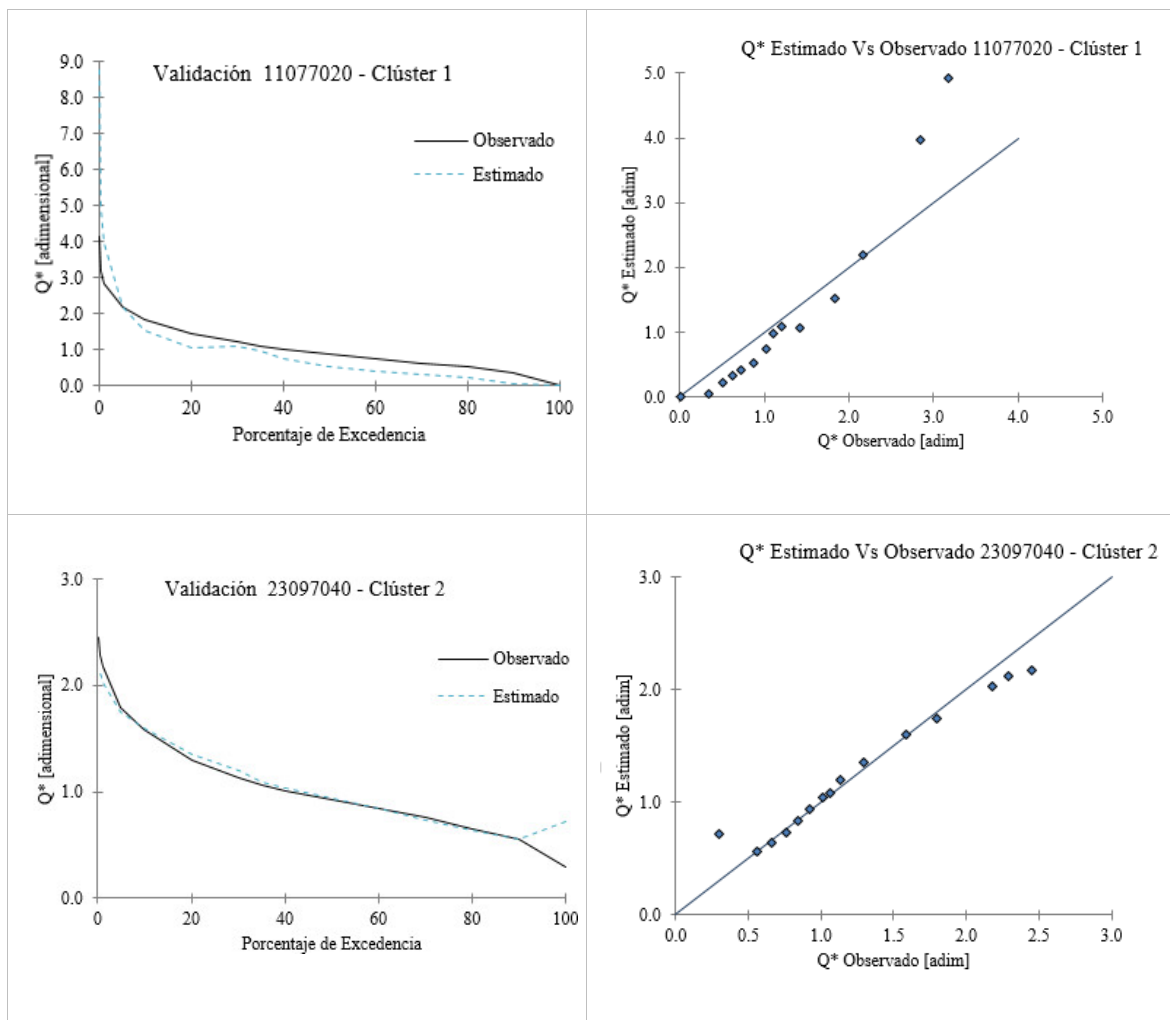
Clúster	R	
	Promedio máximos	Promedio mínimos
11	0.405	-0.440
12	0.286	-0.284
13	0.269	-0.241
14	0.225	-0.210
15	0.258	-0.279

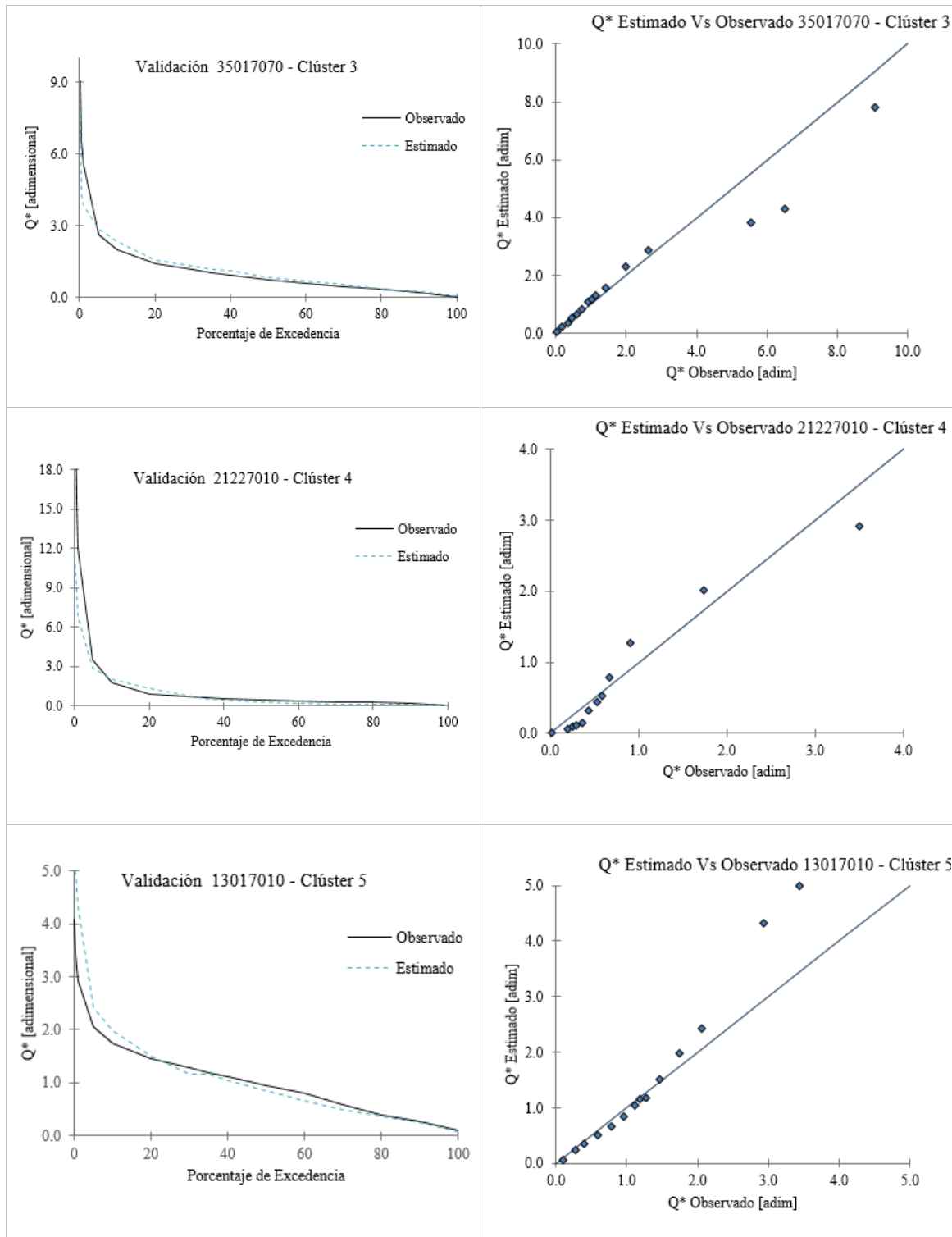
B. Centroides geométricos de cada clúster

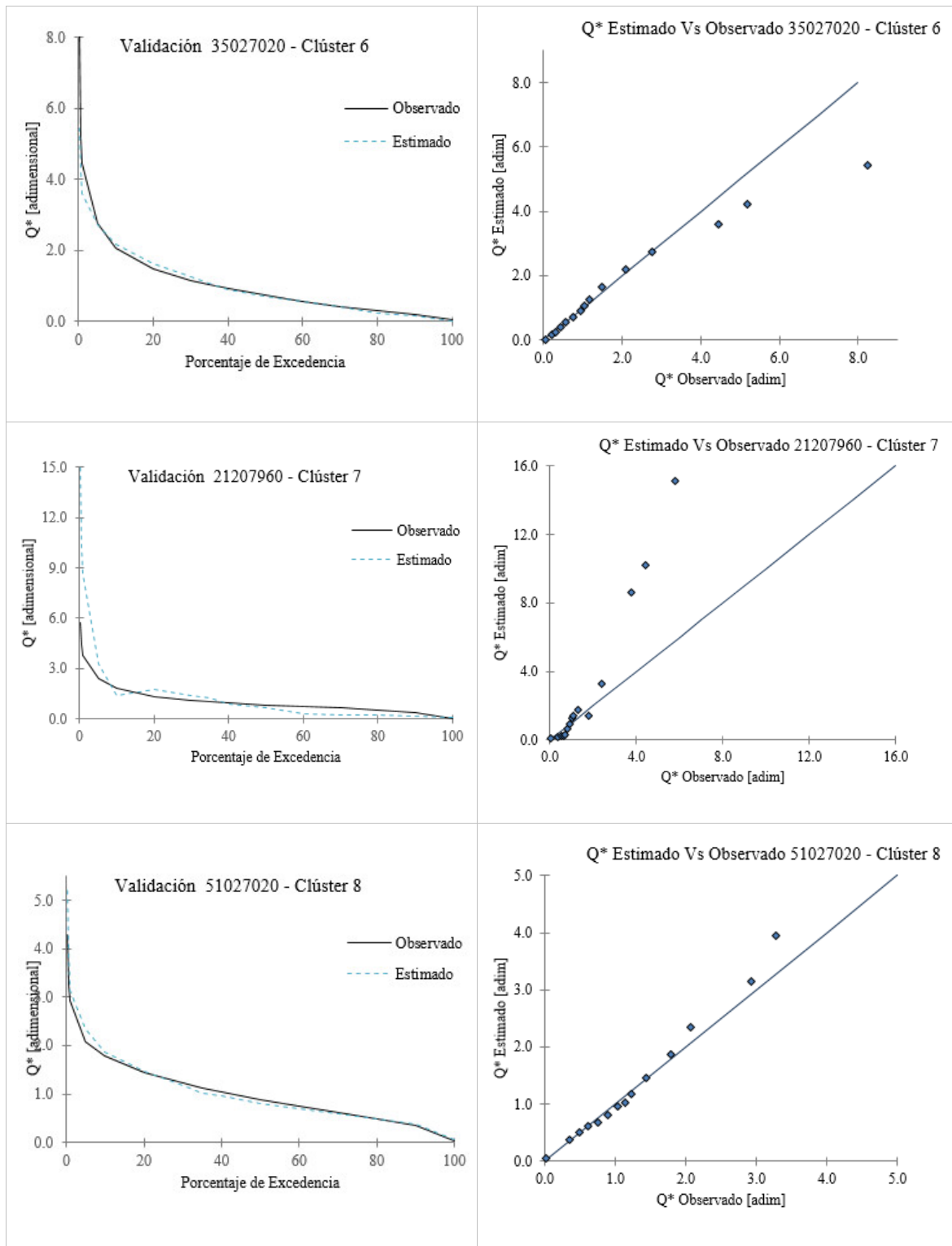
	1	2	3	4	5	6	7
ADre	234.70	88894	464.11	614.92	10157.70	373.85	670.80
Perim	72.11	2300.95	119.09	102.66	586.59	81.25	114.88
Komp	1.46	2.36	1.87	1.56	1.81	1.38	1.50
%Agr	8.51	10.60	1.55	8.47	7.37	52.91	16.20
%Bosq	13.77	15.12	13.83	22.17	41.43	14.54	10.52
%Urb	0.15	1.07	0.03	0.18	0.18	0.13	6.28
DenF	4.88	66.83	3.93	17.62	11.07	7.97	10.30
DenD	20.56	25.32	24.50	23.16	26.26	19.71	21.10
%Lim	0.00	0.00	0.43	0.00	0.01	0.00	0.00
%Aren	2.27	9.82	78.12	10.73	19.02	1.55	2.93
%Arc	0.50	0.18	0.09	1.50	2.01	0.00	0.00
ETPa	752.39	757.87	976.29	957.56	898.26	661.82	660.59
Emax	2725.30	5338.10	580.75	967.72	3947.53	3377.66	3236.63
Emed	1540.04	1567.71	210.91	313.24	680.89	2195.50	2232.42
Emin	934.87	158.52	157.08	95.93	95.42	1302.45	1568.00
Desn	1790.43	5179.57	423.67	871.79	3852.11	2075.21	1668.63
LonC	30.72	957.14	54.38	43.74	233.70	31.65	45.47
Pend	28.78	25.92	3.05	15.03	17.95	31.54	19.73
Pmáx	461.19	446.57	603.62	511.35	557.53	436.79	417.18
Pmed	168.08	162.71	209.33	158.86	188.62	142.47	156.95
Tmed	19.99	19.77	26.53	26.21	24.09	16.04	15.78
H10	41.61	52.49	23.89	44.62	41.19	44.41	31.48
H25	31.01	38.96	16.72	31.05	21.18	35.25	24.91
H50	20.45	23.07	11.79	19.23	6.69	24.84	16.97
H75	11.78	10.95	7.92	10.48	2.68	16.21	11.48

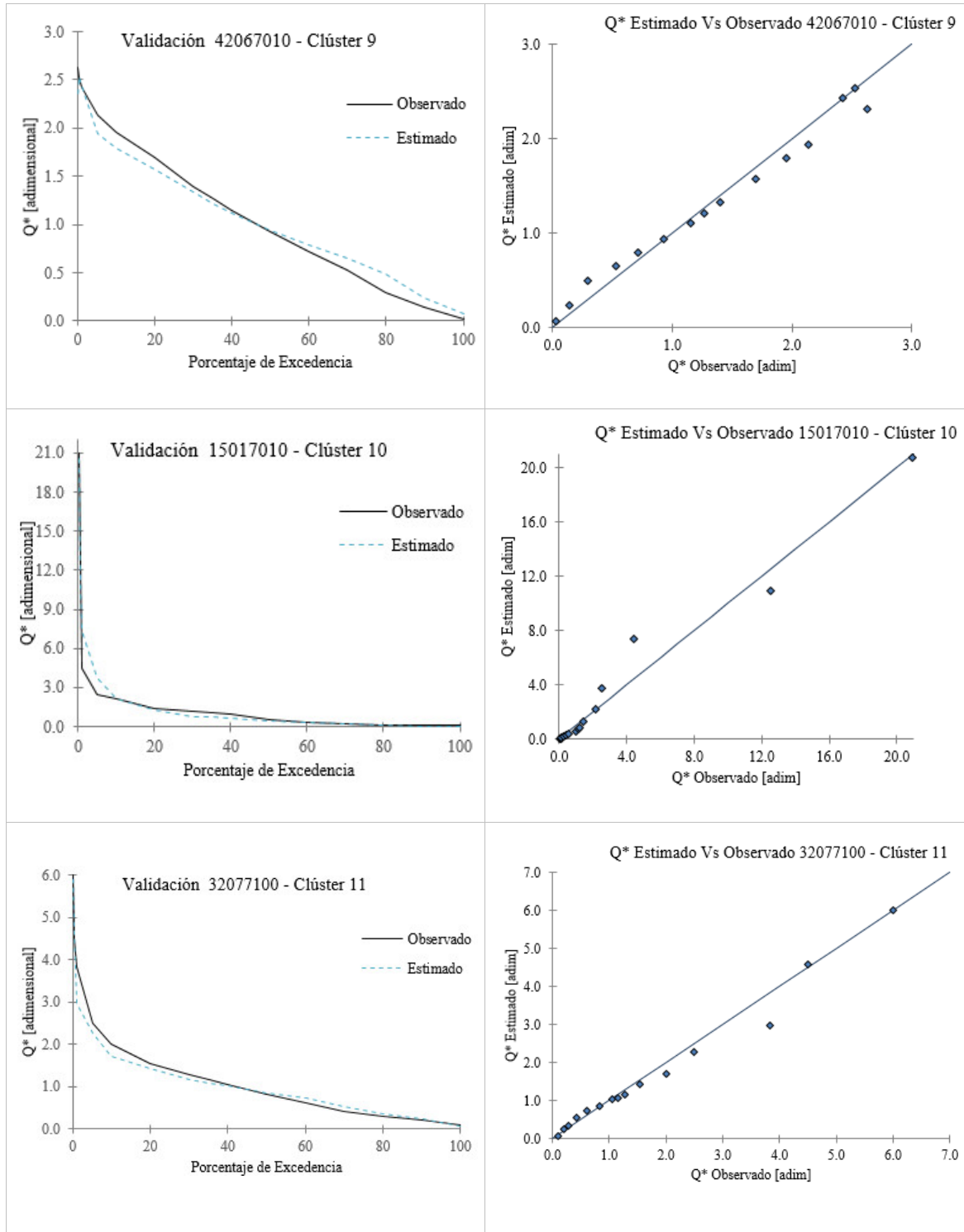
	8	9	10	11	12	13	14	15
ADre	1250.37	60581.05	3306.88	3494.43	1071.95	451.90	231.40	4124.32
Perim	166.18	2042.21	281.80	285.49	162.58	96.70	60.79	331.44
Komp	1.48	2.40	1.68	1.68	1.55	1.43	1.44	1.73
%Agr	12.87	1.18	14.81	5.16	3.44	8.10	16.45	16.42
%Bosq	24.88	72.37	17.32	61.57	61.87	31.03	6.15	20.95
%Urb	0.07	0.07	0.41	0.02	0.08	0.14	0.20	0.15
DenF	9.31	6.15	21.13	13.39	23.41	6.19	8.27	23.72
DenD	22.29	25.90	25.59	26.08	24.16	22.28	18.72	21.55
%Lim	0.00	0.43	0.04	0.00	0.00	0.00	0.00	0.00
%Aren	0.91	15.48	3.41	8.25	1.74	0.60	2.01	1.48
%Arc	0.16	17.23	0.12	0.77	0.39	0.14	0.00	0.04
ETPa	662.80	951.23	754.91	888.59	739.90	587.69	553.98	617.56
Emax	3926.63	2469.31	3933.40	3029.89	3786.93	4031.88	3706.97	4395.85
Emed	2211.64	364.58	1564.99	724.40	1676.81	2788.07	3056.77	2581.46
Emin	687.55	95.54	434.32	113.95	329.07	1660.20	2607.68	521.80
Desn	3239.08	2373.77	3499.09	2915.95	3457.86	2371.68	1099.29	3874.05
LonC	69.10	896.30	125.54	124.07	76.34	39.15	23.79	133.48
Pend	38.58	6.23	31.66	23.99	39.66	33.86	19.35	34.02
Pmáx	418.90	628.63	506.10	1038.81	661.38	400.20	287.22	444.10
Pmed	143.43	248.12	154.52	423.26	241.22	141.72	92.74	146.99
Tmed	15.92	25.34	19.67	23.99	18.89	12.42	10.78	13.68
H10	62.09	28.55	56.80	46.87	64.06	43.28	20.71	69.48
H25	52.66	14.60	41.84	30.61	50.14	36.49	16.32	62.18
H50	39.24	10.02	26.37	14.19	33.55	27.61	11.60	50.35
H75	25.27	7.43	14.07	5.86	19.41	18.81	7.38	33.41

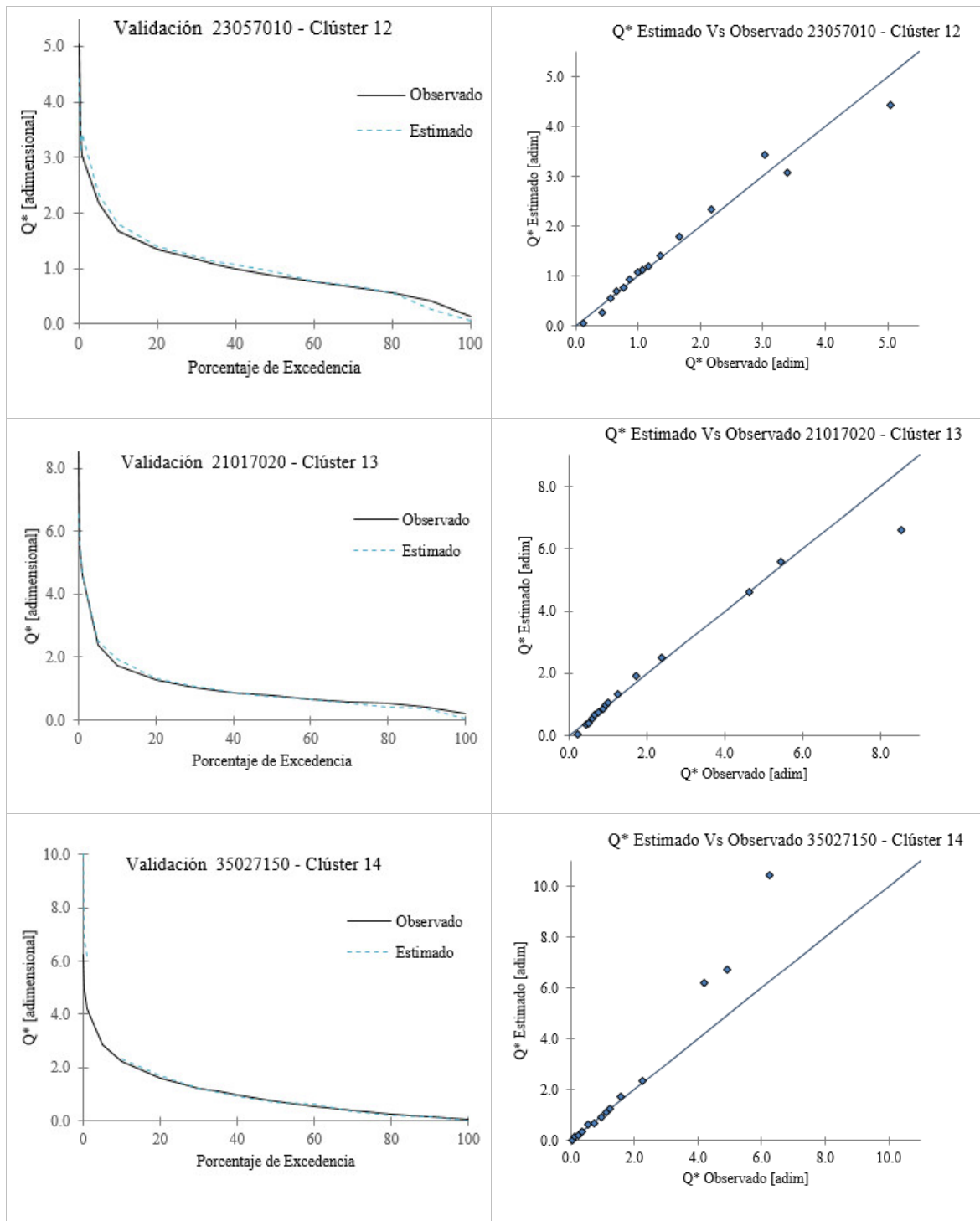
C. Resultados de Regresiones Lineales Múltiples

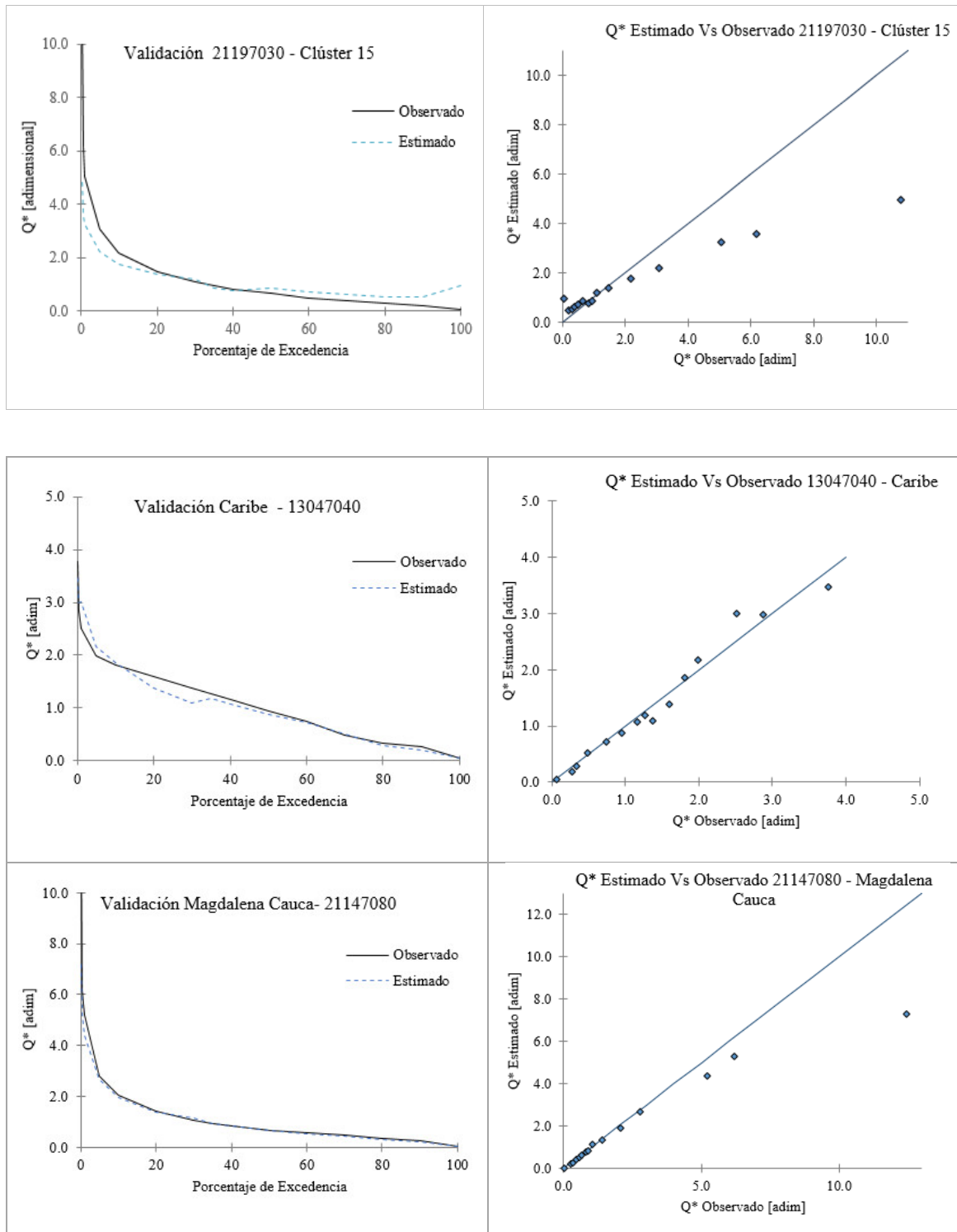


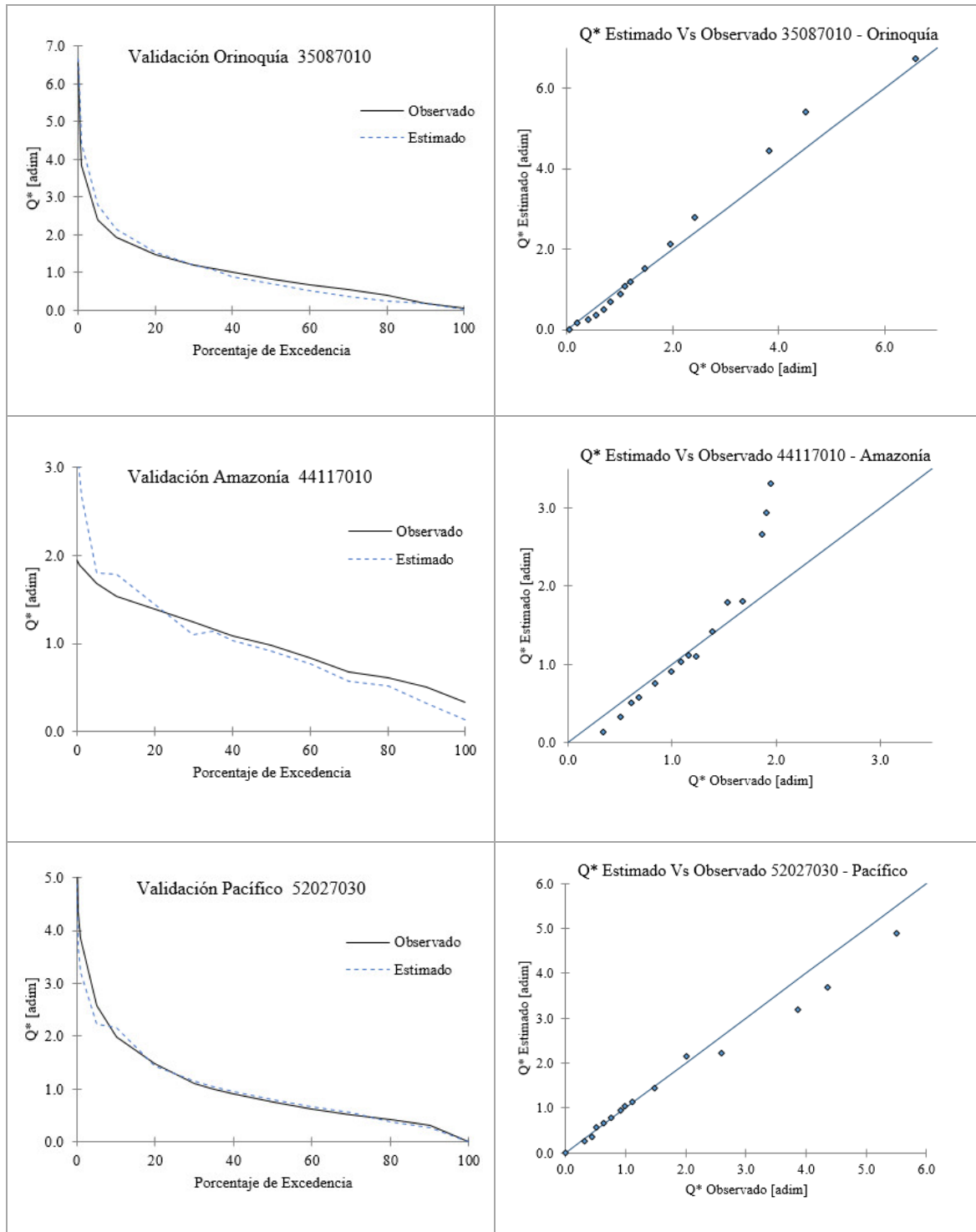












Ecuaciones de Regresión Clúster #1:

Caudal Q_p		X_1	α_1	X_2	α_2	X_3	α_3	X_4	α_4	R^2
100	5.5E-42	ADre	0.534	ETPa	9.874	Emin	1.54	Pmáx	2.174	0.29
90	2.4E-07	%Agr	-0.088	%Urb	0.164	DenF	0.21	Pmáx	2.363	0.46
80	2.3E-05	%Agr	-0.043	%Urb	0.085	LonC	0.46	Pmáx	1.386	0.37
70	4.8E-04	Perim	0.378	%Agr	-0.036	%Urb	0.06	Pmáx	0.907	0.40
60	1.6E-03	Perim	0.342	%Agr	-0.030	%Urb	0.05	Pmáx	0.764	0.44
50	8.3	%Urb	0.037	ETPa	-1.294	LonC	0.27	Pmáx	0.872	0.45
40	8.8E-04	%Agr	-0.022	Emax	0.412	LonC	0.21	Pmáx	0.466	0.46
35	2.5E-02	%Agr	-0.028	%Bosq	0.029	Emax	0.66	Pend	-0.51	0.47
30	6.0E-02	%Agr	-0.023	%Bosq	0.024	Emax	0.52	Pend	-0.41	0.49
20	6.8E-14	DenF	0.025	DenD	-0.139	ETPa	3.47	Emed	1.09	0.38
10	2.1E	Perim	-0.115	DenF	0.032	DenD	-0.07	H10	0.16	0.36
5	0.37	ADre	-0.094	%Agr	0.013	Tmed	0.58	H10	0.18	0.55
1	1.7E-06	%Agr	0.030	DenD	0.107	ETPa	2.35	LonC	-0.31	0.56
0.5	2.2E-07	%Agr	0.039	DenD	0.147	ETPa	2.72	LonC	-0.39	0.54
0.1	6.5E+02	%Agr	0.052	DenD	0.198	Emin	-0.41	LonC	-0.60	0.58

Ecuaciones de Regresión Clúster #2

Caudal Q_p	β	X_1	α_1	X_2	α_2	R^2
100	2.098E+08	%Bosq	-7.017	%Arc	0.459	0.88
90	1.064E-44	ADre	-0.140	Emax	11.902	0.38
80	4.094E-23	Komp	0.202	Emax	5.933	0.59
70	7.204E-06	Desn	1.466	H25	-0.274	0.76
60	5.818E-02	Emed	-0.202	Desn	0.486	0.85
50	6.747E-04	ETPa	1.030	H50	0.132	0.93
40	4.423E-01	%Urb	-0.071	Pmed	0.172	0.70
35	9.567E+07	%Urb	-0.143	Emax	-2.127	0.41
30	6.487E+00	%Agr	-0.090	DenD	-0.459	0.32
20	6.893E+12	%Arc	0.008	Emax	-3.406	0.59
10	3.108E+00	Perim	-0.177	H25	0.183	0.80
5	9.198E+01	Perim	-0.515	%Urb	-0.357	0.70
1	3.475E+00	Pend	0.887	H10	-0.858	0.61
0.5	8.897E+35	ETPa	-9.685	Emed	-2.420	0.55
0.1	2.906E-01	%Agr	0.767	%Arc	-0.083	0.55

Ecuaciones de Regresión Clúster #3

Caudal Q_p	β	X_1	α_1	X_2	α_2	R^2
100	1.39E-05	%Agr	0.265	Pmed	1.516	0.76
90	1.24E-08	Desn	0.614	Pmed	2.406	0.77
80	2.54E-04	Pmed	1.490	H75	-0.898	0.84
70	7.83E-05	Pmed	1.742	H75	-0.750	0.87
60	5.65E-06	Emax	0.524	Pmed	1.502	0.83
50	8.13E-05	Emax	0.438	Pmed	1.164	0.78
40	5.58E+00	Pmáx	-2.554	Pmed	2.695	0.81
35	1.04E+01	Pmáx	-2.319	Pmed	2.327	0.86
30	2.20E+01	Pmáx	-2.127	Pmed	1.989	0.89
20	2.84E+01	Pmáx	-1.294	Pmed	1.007	0.81
10	1.09E+00	H50	0.968	H75	-0.756	0.70
5	1.93E+00	%Bosq	-0.033	H50	0.227	0.68
1	4.83E+00	%lim	-0.101	Pend	-0.698	0.69
0.5	5.41E+00	%lim	-0.120	Pend	-0.802	0.71
0.1	6.71E+00	%lim	-0.135	Pend	-0.946	0.74

Ecuaciones de Regresión Clúster #4

Q_p	β	X_1	α_1	X_2	α_2	X_3	α_3	X_4	α_4	X_5	α_5	R^2
100	1.4E+115	Komp	-5.658	ETPa	-38.941	Emed	-2.267	LonC	0.695	Pmed	1.935	0.31
90	8.8E+69	%Bosq	0.116	ETPa	-23.183	Emed	-1.137	Pmáx	-0.926	Pmed	1.507	0.37
80	3.1E+73	%Agr	0.031	%Bosq	0.106	ETPa	-24.666	Emed	-1.277	Pmed	1.025	0.43
70	7.5E+88	%Bosq	0.111	%Aren	0.043	ETPa	-29.496	Emed	-1.491	Pmed	0.884	0.45
60	8.7E+80	%Bosq	0.107	%Aren	0.046	ETPa	-26.786	Emed	-1.390	Pmed	0.782	0.44
50	8.8E+67	%Bosq	0.084	ETPa	-22.549	Emed	-1.247	LonC	0.244	Pmed	0.687	0.40
40	3.8E+64	%Bosq	0.077	ETPa	-21.359	Emed	-1.195	LonC	0.264	Pmed	0.589	0.38
35	2.3E+63	%Bosq	0.077	ETPa	-20.919	Emed	-1.168	LonC	0.265	Pmed	0.548	0.37
30	1.3E+31	%Bosq	0.029	ETPa	-10.319	Emed	-0.609	LonC	0.138	Pmed	0.384	0.50
20	8.8E+10	DenD	0.154	%Arc	0.024	ETPa	-3.648	Emed	-0.246	Pmed	0.228	0.36
10	2.1E-01	%Agr	-0.009	%Urb	-0.012	DenD	0.190	%Aren	-0.007	Tmed	0.501	0.23
5	2.9E+00	ADre	-0.053	%Urb	-0.028	DenD	0.214	Emed	0.076	Pmed	-0.183	0.31
1	8.4E-34	Komp	1.375	ETPa	11.216	Emed	0.837	LonC	-0.360	Pmed	-0.591	0.46
0.5	1.3E-41	Komp	1.595	ETPa	13.830	Emed	1.036	LonC	-0.416	Pmed	-0.730	0.50
0.1	4.7E-47	ADre	-0.157	ETPa	15.024	Emed	1.146	Pmáx	1.470	Pmed	-1.719	0.53

Ecuaciones de Regresión Clúster #5

Caudal Q_p	β	X_1	α_1	X_2	α_2	R^2
100	5.47E-05	Komp	3.94	%Bosq	1.197	0.59
90	3.44E-02	%Bosq	0.57	%Urb	0.087	0.64
80	6.90E-02	%Bosq	0.50	%Urb	0.087	0.65
70	1.17E-01	%Bosq	0.47	%Urb	0.093	0.69
60	1.95E-01	%Bosq	0.41	%Urb	0.086	0.70
50	1.02E-01	Emed	-0.35	Pmed	0.831	0.67
40	2.23E-01	Emed	-0.28	Pmed	0.629	0.61
35	3.33E-01	Emed	-0.23	Pmed	0.519	0.55
30	1.64E-01	LonC	0.15	Pmed	0.228	0.51
20	3.30	Komp	-0.46	Pend	-0.172	0.34
10	2.92	%Bosq	-0.14	%Urb	-0.033	0.45
5	4.91	%Bosq	-0.27	%Urb	-0.071	0.58
1	1.11E-05	Desn	2.00	LonC	-0.729	0.65
0.5	1.41E-06	Desn	2.34	LonC	-0.852	0.67
0.1	2.30E-07	Desn	2.70	LonC	-1.046	0.67

Ecuaciones de Regresión Clúster #6

Q_p	β	X_1	α_1	X_2	α_2	X_3	α_3	X_4	α_4	X_5	α_5	R^2
100	4.50E+280	%Agr	-2.56	DenF	0.24	DenD	-3.21	ETPa	-62.09	Emed	-29.70	0.48
90	1.70E+14	DenD	0.72	Emin	-2.62	Pend	1.18	Tmed	-3.01	H10	-3.56	0.47
80	1.25E+12	DenD	0.49	Emin	-2.20	Pend	1.01	Tmed	-2.51	H10	-2.96	0.45
70	3.20E+09	%Agr	0.29	Emin	-1.71	Pend	0.70	Tmed	-2.35	H10	-2.00	0.48
60	3.59E+07	%Agr	0.20	Emin	-1.34	Pend	0.51	Tmed	-1.98	H10	-1.45	0.51
50	1.19E+06	%Agr	0.16	Emin	-1.07	Pend	0.40	Tmed	-1.67	H10	-1.09	0.50
40	8.47E+05	Emin	-0.94	Desn	-0.21	Pend	0.35	Tmed	-1.58	H10	-0.63	0.50
35	2.10E+04	%Urb	0.03	Emin	-0.60	Desn	-0.44	Pend	0.37	Tmed	-1.28	0.52
30	2.07E+04	%Urb	0.03	Emin	-0.58	Desn	-0.43	Pend	0.35	Tmed	-1.29	0.48
20	3.51E+00	Komp	0.37	DenD	-0.30	Pmáx	-0.10	Tmed	-0.23	H25	0.31	0.49
10	9.83E+18	ETPa	-8.48	Pend	-0.23	Tmed	4.00	H10	0.55	H75	-0.13	0.52
5	2.16E-09	Emax	0.55	Emin	1.16	Pend	-0.51	Tmed	2.06	H10	1.12	0.64
1	9.78E-11	%Urb	-0.07	Emin	1.37	Desn	1.29	Pend	-1.07	Tmed	2.98	0.57
0.5	1.28E+00	%Bosq	0.09	%Urb	-0.06	DenD	0.44	Pend	-1.16	Tmed	1.34	0.60
0.1	5.90E-24	%Urb	-0.08	DenD	0.73	Emed	5.03	Pend	-1.27	Tmed	6.79	0.69

Ecuaciones de Regresión Clúster #7

Caudal Q_p	β	X_1	α_1	X_2	α_2	R^2
100	1.47E+183	ETPa	-85.262	Tmed	46.733	0.69
90	4.25E-02	Pend	1.262	H25	-0.581	0.61
80	6.51E-02	Pend	1.111	H25	-0.480	0.71
70	7.67E-02	Pend	0.915	H50	-0.338	0.76
60	8.69E-02	Pend	0.913	H50	-0.318	0.75
50	3.89E-04	Emed	0.644	Pmáx	0.436	0.78
40	1.31E-02	%Agr	-0.024	Emax	0.518	0.74
35	3.12E-02	%Aren	0.033	Emed	0.463	0.75
30	3.32E-02	%Aren	0.035	Emed	0.469	0.78
20	6.17E-02	%Aren	0.034	Emed	0.417	0.76
10	3.09E-01	Desn	0.315	H75	-0.266	0.65
5	2.00E+01	Emin	-0.166	Pend	-0.298	0.64
1	1.12E+03	Emin	-0.443	Pend	-0.806	0.70
0.5	1.41E+03	Emin	-0.445	Pend	-0.824	0.68
0.1	3.53E+03	Emin	-0.488	Pend	-0.921	0.58

Ecuaciones de Regresión Clúster #8

Caudal Q_p	β	X_1	α_1	X_2	α_2	X_3	α_3	X_4	α_4	R^2
100	1.5E-03	ADre	-1.055	%Bosq	0.352	DenF	-0.297	LonC	2.491	0.42
90	5.5E+00	Komp	2.044	%Bosq	0.140	H10	-4.044	H25	3.195	0.45
80	3.0	Komp	1.289	%Bosq	0.114	H10	-2.655	H25	2.070	0.47
70	2.2	Komp	0.812	%Bosq	0.098	H10	-1.843	H25	1.420	0.51
60	2.4E-01	%Bosq	0.091	%Urb	0.013	DenD	0.718	Pmáx	-0.231	0.55
50	3.5E-01	%Bosq	0.077	%Urb	0.012	DenD	0.547	Pmáx	-0.172	0.54
40	3.7E-01	%Agr	-0.005	%Bosq	0.057	%Urb	0.011	DenD	0.261	0.57
35	9.3E-01	%Agr	-0.006	%Bosq	0.043	%Urb	0.009	DenF	0.012	0.55
30	7.8E-01	%Agr	-0.005	%Bosq	0.030	DenF	0.020	Emin	0.037	0.58
20	9.5E-01	Komp	-0.234	DenF	0.018	Pend	-0.112	Pmáx	0.140	0.32
10	2.3	%Bosq	-0.026	DenD	-0.345	Pend	-0.134	Pmáx	0.236	0.36
5	1.5	ADre	0.043	%Bosq	-0.039	%Urb	-0.017	%Arc	-0.026	0.34
1	5.1E+02	%Bosq	-0.054	%Arc	-0.022	Emed	-0.414	Pmed	-0.344	0.49
0.5	3.7E-01	Komp	-0.745	%Bosq	-0.103	H10	2.569	H25	-1.897	0.51
0.1	5.3E+06	%Aren	-0.044	Emed	-1.087	Pend	-0.595	Pmed	-0.650	0.59

Ecuaciones de Regresión Clúster #9

Caudal Q_p	β	X_1	α_1	X_2	α_2	X_3	α_3	R^2
100	2.5E-14	Komp	3.040	DenD	9.673	Emed	-0.886	0.42
90	7.0E-02	Komp	1.212	%Urb	-0.063	%Arc	0.034	0.62
80	1.6E-01	Komp	0.727	%Urb	-0.129	H75	-0.156	0.63
70	2.2E-05	DenF	-0.049	Desn	0.182	Pmed	1.603	0.65
60	5.1E-01	%Bosq	0.114	DenF	-0.015	H75	-0.059	0.51
50	1.5	%Agr	0.018	Emed	-0.117	Emin	0.058	0.44
40	2.3E+01	Desn	-0.065	Pmed	-0.451	H50	-0.038	0.77
35	1.5E+01	%Arc	-0.002	Desn	-0.031	Pmed	-0.420	0.70
30	8.8E+01	Komp	-0.170	DenD	-0.721	Pmed	-0.308	0.49
20	1.8E+01	Komp	-0.225	%Urb	0.015	DenD	-0.676	0.35
10	2.4	Komp	-0.377	DenF	0.018	H75	0.042	0.41
5	2.7	Komp	-0.463	DenF	0.023	H75	0.052	0.46
1	5.6	%Agr	-0.028	%Bosq	-0.244	DenF	0.024	0.55
0.5	6.1	%Agr	-0.031	%Bosq	-0.257	DenF	0.025	0.54
0.1	3.7	%Agr	-0.059	%Bosq	-0.366	Emax	0.122	0.51

Ecuaciones de Regresión Clúster #10

Caudal Q_p	β	X_1	α_1	X_2	α_2	X_3	α_3	X_4	α_4	R^2
100	1.9E-08	Emed	3.131	Desn	-2.362	LonC	1.768	H75	1.048	0.50
90	1.3E-08	%Lim	-0.289	Emed	1.194	LonC	0.323	Pmed	0.918	0.61
80	2.4E+33	%Lim	-0.099	ETPa	-15.908	LonC	0.237	Tmed	8.644	0.62
70	3.2E-06	%Lim	-0.183	Emed	0.856	LonC	0.208	Pmed	0.674	0.63
60	2.7E-05	%Lim	-0.140	Emed	0.716	LonC	0.164	Pmed	0.597	0.60
50	2.6E+14	Perim	0.103	ETPa	-7.088	Pmed	0.374	Tmed	3.681	0.56
40	5.3E-06	Perim	0.057	Emed	0.876	Pmed	0.308	Tmed	1.250	0.48
35	3.5E+02	Perim	0.071	%Bosq	0.019	ETPa	-1.241	Pmáx	0.306	0.42
30	7.0E-02	Perim	0.077	%Bosq	0.018	Emin	0.080	Pmáx	0.295	0.37
20	1.0E+00	Komp	0.238	Pmáx	0.141	H25	-0.154	H75	-0.041	0.30
10	1.8E+01	%Bosq	-0.018	Emed	-0.238	LonC	-0.046	H75	-0.083	0.43
5	8.6E+02	Emed	-0.422	LonC	-0.099	Pmed	-0.317	H50	-0.199	0.38
1	1.4E-07	%Bosq	-0.040	%Arc	0.046	ETPa	2.822	LonC	-0.238	0.44
0.5	4.2E-07	%Bosq	-0.045	ETPa	3.164	LonC	-0.305	Pmáx	-0.510	0.47
0.1	1.6E-05	%Urb	0.027	ETPa	2.808	LonC	-0.510	Pmed	-0.628	0.51

Ecuaciones de Regresión Clúster #11

Caudal Q_p	β	X_1	α_1	X_2	α_2	R^2
100	7.74E-07	%Bosq	1.560	E _{max}	0.593	0.63
90	1.01E-03	LonC	0.436	P _{med}	0.629	0.67
80	4.69E-03	LonC	0.368	P _{med}	0.478	0.70
70	1.01E-01	%Bosq	0.156	LonC	0.242	0.72
60	2.54E-01	%Arc	0.017	LonC	0.245	0.76
50	4.23E-01	DenF	-0.055	LonC	0.175	0.76
40	6.26E-01	Perim	0.107	DenF	-0.060	0.68
35	6.55E-01	%Bosq	0.106	%Aren	0.028	0.63
30	8.07E-01	%Bosq	0.077	%Aren	0.025	0.58
20	1.47E+00	%Bosq	0.060	LonC	-0.065	0.43
10	3.36E+00	%Arc	-0.013	LonC	-0.155	0.76
5	6.46E+00	%Aren	-0.025	LonC	-0.229	0.83
1	2.19E+01	DenF	0.128	LonC	-0.495	0.85
0.5	2.18E+07	ETPa	-1.938	LonC	-0.541	0.85
0.1	9.14E+09	ETPa	-2.721	LonC	-0.656	0.85

Ecuaciones de Regresión Clúster #12

Caudal Q_p	β	X_1	α_1	X_2	α_2	X_3	α_3	X_4	α_4	R^2
100	1.4E-02	Adre	0.445	%Agr	-0.157	DenD	5.944	H10	-4.932	0.44
90	5.2E+01	ADre	0.135	%Agr	-0.048	H10	-1.712	H75	0.369	0.37
80	1.3E+01	Perim	0.209	%Agr	-0.027	H10	-1.325	H75	0.342	0.37
70	3.8E+00	Perim	0.225	DenF	-0.068	H25	-1.025	H75	0.393	0.43
60	1.5E+00	Perim	0.167	DenF	-0.062	H50	-0.928	H75	0.620	0.49
50	1.8E+12	DenF	-0.035	ETPa	-6.332	Desn	0.254	T _{med}	3.881	0.44
40	1.3E-07	DenF	-0.027	DenD	0.420	E _{med}	1.172	T _{med}	2.001	0.42
35	9.8E-02	ADre	0.049	DenF	-0.018	DenD	0.580	T _{med}	0.044	0.41
30	1.2E-01	ADre	0.038	DenF	-0.014	DenD	0.579	E _{min}	0.038	0.39
20	4.3E-01	%Aren	-0.013	P _{máx}	0.084	H25	0.262	H75	-0.148	0.34
10	3.6E-09	%Agr	0.008	%Aren	-0.008	ETPa	4.133	T _{med}	-2.477	0.39
5	2.5E-12	Perim	-0.049	%Agr	0.015	ETPa	5.759	T _{med}	-3.479	0.38
1	8.5E+00	ADre	-0.193	DenF	0.091	DenD	-1.500	H10	1.257	0.51
0.5	2.0E+00	DenF	0.107	LonC	-0.520	P _{end}	-0.736	H10	1.293	0.56
0.1	5.4E+01	Perim	-0.553	%Urb	-0.059	DenF	0.135	%Aren	0.059	0.57

Ecuaciones de Regresión Clúster #13

Q_p	β	X_1	α_1	X_2	α_2	X_3	α_3	X_4	α_4	X_5	α_5	R^2
100	2.4E-03	%Agr	-0.15	E _{max}	8.61	E _{med}	-15.68	E _{min}	6.44	H75	2.88	0.31
90	1.2E-03	Komp	0.87	%Agr	-0.06	DenF	0.07	DenD	1.02	Pend	0.58	0.24
80	1.3E-03	Perim	-0.19	%Agr	-0.03	Desn	0.47	Pmed	0.44	H75	0.29	0.42
70	1.2E-02	%Agr	-0.02	DenD	0.21	Pend	0.31	Pmed	0.31	H75	0.19	0.40
60	3.5E-02	%Agr	-0.01	DenD	0.14	Pend	0.23	Pmed	0.24	H75	0.17	0.41
50	4.6E-02	%Agr	-0.01	E _{max}	0.37	E _{med}	-0.28	P _{máx}	0.24	H75	0.17	0.37
40	1.2E-01	%Agr	-0.01	%Aren	0.01	Desn	0.10	P _{máx}	0.17	H75	0.08	0.32
35	1.6E-01	Komp	-0.11	Desn	0.07	LonC	0.04	P _{máx}	0.17	H75	0.06	0.28
30	2.2	DenF	0.01	%Aren	0.01	E _{min}	-0.14	P _{máx}	0.12	Tmed	-0.15	0.33
20	1.2	Komp	-0.10	DenF	0.01	%Aren	0.01	%Arc	-0.01	LonC	0.03	0.38
10	2.1	Perim	0.11	%Agr	0.01	E _{max}	-0.54	E _{med}	0.57	H75	-0.20	0.44
5	7.2E+01	Komp	0.33	%Agr	0.02	Desn	-0.20	Pmed	-0.28	H75	-0.20	0.34
1	1.0E+03	%Agr	0.03	%Aren	-0.02	Desn	-0.23	P _{máx}	-0.52	H75	-0.25	0.35
0.5	2.1E+03	%Agr	0.03	%Aren	-0.02	Desn	-0.24	P _{máx}	-0.59	H75	-0.29	0.36
0.1	1.9E+03	LonC	-0.25	Pend	-0.50	Pmed	-0.87	H25	0.79	H75	-0.55	0.36

Ecuaciones de Regresión Clúster #14

Q_p	β	X_1	α_1	X_2	α_2	X_3	α_3	X_4	α_4	X_5	α_5	R^2
100	9.5E-20	Komp	2.20	%Urb	0.25	E _{min}	3.19	Pend	1.05	Pmed	2.52	0.27
90	5.2E+67	Perim	0.37	DenF	-0.05	ETPa	-29.74	Pmed	1.57	Tmed	9.17	0.46
80	3.9E+52	Perim	0.24	DenF	-0.04	ETPa	-23.16	Pmed	1.25	Tmed	7.32	0.46
70	3.0E-08	Komp	0.88	%Bosq	0.02	E _{min}	1.23	Desn	0.27	P _{máx}	0.80	0.45
60	2.1E-06	Perim	0.11	%Bosq	0.02	DenF	-0.02	E _{med}	1.14	Pmed	0.63	0.43
50	7.7E+15	%Bosq	0.01	ETPa	-7.09	Desn	0.07	Pmed	0.50	Tmed	2.10	0.36
40	1.0E-03	%Bosq	0.01	E _{med}	0.62	Desn	0.07	Pend	-0.11	Pmed	0.34	0.33
35	6.5E+06	%Bosq	0.01	ETPa	-1.90	E _{min}	-0.59	Pend	-0.15	Pmed	0.27	0.31
30	6.0E+07	%Bosq	0.01	ETPa	-1.99	E _{min}	-0.71	Pend	-0.19	Pmed	0.19	0.28
20	5.5E+05	Komp	-0.30	DenF	0.01	E _{min}	-1.28	Pend	-0.31	Tmed	-0.79	0.33
10	1.6E-16	Perim	-0.05	%Aren	0.01	ETPa	6.93	P _{máx}	-0.21	Tmed	-2.23	0.28
1	1.1E+13	Komp	-0.66	%Bosq	-0.03	E _{med}	-2.73	P _{máx}	-0.60	Tmed	-1.22	0.37
0.5	3.2E+05	Komp	-0.78	%Bosq	-0.03	DenF	0.01	E _{med}	-0.84	P _{máx}	-0.68	0.34
0.1	6.6E+05	Komp	-1.05	%Bosq	-0.03	DenF	0.02	E _{med}	-0.81	P _{máx}	-0.77	0.29

Ecuaciones de Regresión Clúster #15

Q_p	β	X_1	α_1	X_2	α_2	X_3	α_3	X_4	α_4	X_5	α_5	R^2
100	2.2E-15	Perim	4.72	%Urb	-0.51	%Are	-0.19	LonC	-2.90	H75	4.47	0.38
90	4.7E-08	ADre	0.37	%Urb	-0.14	DenD	3.42	%Are	-0.04	Emin	0.27	0.60
80	3.5E-08	ADre	0.25	%Urb	-0.08	DenD	2.58	Emed	0.64	Emin	0.17	0.63
70	8.9E+23	ADre	0.19	%Urb	-0.08	DenD	1.28	ETPa	-11.45	Tmed	4.61	0.56
60	5.0E+22	ADre	0.16	%Urb	-0.07	DenD	0.54	ETPa	-10.48	Tmed	4.37	0.50
50	3.1E+01	ADre	0.06	%Urb	-0.06	%Arc	0.04	Pend	-0.37	H25	-0.69	0.49
40	1.4E-16	%Urb	-0.02	%Lim	0.10	Emed	3.89	Tmed	3.66	H10	-0.71	0.52
35	4.1E-15	%Urb	-0.02	%Lim	0.09	Emed	3.52	Tmed	3.22	H25	-0.56	0.51
30	3.3E+01	%Urb	-0.03	%Arc	0.03	Pend	-0.29	H25	-0.43	H75	-0.15	0.47
20	1.5E+01	Komp	-0.16	DenD	-0.49	H25	-0.31	H50	0.36	H75	-0.26	0.60
10	4.7E+01	ADre	-0.06	%Urb	0.02	DenD	-0.60	Emin	-0.05	H75	-0.14	0.51
5	1.1E-01	%Urb	0.05	DenD	-0.64	%Arc	-0.05	Pend	0.43	H25	0.83	0.34
1	7.9E-27	ADre	-0.20	%Urb	0.09	DenF	0.09	ETPa	11.97	Tmed	-5.30	0.39
0.5	8.2E-28	ADre	-0.24	%Urb	0.10	DenF	0.10	ETPa	12.47	Tmed	-5.47	0.41
0.1	5.9E+04	ADre	-0.32	%Urb	0.10	DenD	-1.72	%Arc	-0.04	Emin	-0.21	0.48

Ecuaciones de Regresión Región Caribe

Caudal Q_p	β	X_1	α_1	X_2	α_2	R^2
100	6.6E+11	ADre	0.640	ETPa	-5.242	0.30
90	6.4E+06	ETPa	-3.705	Pmed	1.486	0.41
80	2.5E+05	ETPa	-2.932	Pmed	1.184	0.38
70	8.5E-03	%Bosq	0.106	Pmed	0.685	0.32
60	3.8E-01	%Bosq	0.191	DenF	-0.123	0.30
50	5.0E-01	%Bosq	0.178	DenF	-0.125	0.28
40	6.4E-01	%Bosq	0.165	DenF	-0.123	0.25
35	7.3E-01	%Bosq	0.160	DenF	-0.124	0.24
30	4.3E+00	Pmáx	-0.866	Pmed	0.781	0.28
20	3.9E+00	Pmáx	-0.477	Pmed	0.377	0.15
10	3.4E+00	H10	-0.500	H25	0.398	0.15
5	1.0E+01	ADre	-0.062	Pmed	-0.196	0.33
1	9.5E+01	ADre	-0.136	Pmed	-0.433	0.40
0.5	2.3E+01	ADre	-0.257	DenF	0.080	0.41
0.1	4.9E+01	ADre	-0.336	DenF	0.117	0.43

Ecuaciones de Regresión Región Magdalena Cauca

Caudal Q_p	β	X_1	α_1	X_2	α_2	R^2
100	1.5E-14	E _{max}	1.548	P _{med}	3.153	0.28
90	8.3E-06	E _{max}	0.698	P _{med}	0.941	0.24
80	7.5E-05	E _{max}	0.589	P _{med}	0.751	0.39
70	3.8E-04	E _{max}	0.502	P _{med}	0.617	0.43
60	1.4E-03	E _{max}	0.432	P _{med}	0.506	0.45
50	4.8E-03	E _{max}	0.362	P _{med}	0.416	0.43
40	2.5E-02	E _{max}	0.248	P _{med}	0.304	0.37
35	5.9E-02	E _{max}	0.191	P _{med}	0.250	0.32
30	4.9E-01	%Bosq	0.017	E _{max}	0.095	0.27
20	1.3E+00	%Bosq	0.005	DenF	0.008	0.10
10	1.1E+01	E _{max}	-0.116	P _{med}	-0.170	0.24
5	6.1E+01	E _{max}	-0.202	P _{med}	-0.313	0.35
1	5.6E+01	AD _{re}	-0.095	P _{med}	-0.404	0.34
0.5	9.8E+01	AD _{re}	-0.114	P _{med}	-0.457	0.35
0.1	2.3E+02	AD _{re}	-0.155	P _{med}	-0.503	0.37

Ecuaciones de Regresión Región Orinoquía

Caudal Q_p	β	X_1	α_1	X_2	α_2	R^2
100	5.001E-04	%Aren	-0.087	LonC	0.840	0.25
90	3.582E-02	%Arc	0.021	E _{max}	0.209	0.16
80	8.139E-02	%Arc	0.021	E _{max}	0.161	0.16
70	1.717E-01	%Arc	0.026	E _{max}	0.119	0.21
60	4.235E-01	%Arc	0.028	Desn	0.054	0.30
50	6.143E-01	AD _{re}	0.035	%Arc	0.015	0.41
40	7.144E-01	Perim	0.060	%Arc	0.011	0.48
35	7.786E-01	AD _{re}	0.044	DenF	-0.011	0.53
30	9.249E-01	AD _{re}	0.037	DenF	-0.010	0.52
20	1.315E+00	AD _{re}	0.022	DenF	-0.009	0.37
10	2.628E+00	%Arc	-0.005	P _{med}	-0.050	0.15
5	3.167E+00	%Arc	-0.013	LonC	-0.053	0.51
1	7.765E+00	%Arc	-0.023	LonC	-0.168	0.69
0.5	1.339E+01	Perim	-0.216	%Arc	-0.026	0.71
0.1	1.857E+01	%Arc	-0.030	LonC	-0.282	0.71

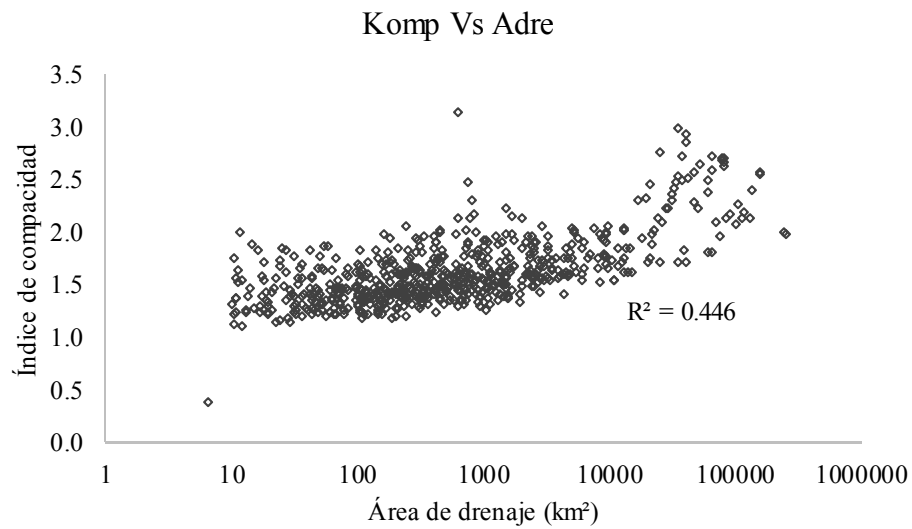
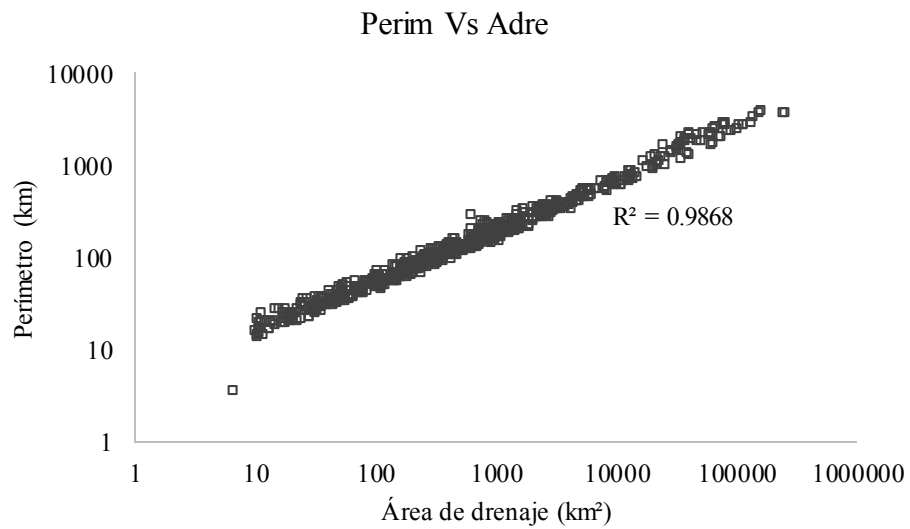
Ecuaciones de Regresión Región Amazonía

Caudal Q_p	β	X_1	α_1	X_2	α_2	R^2
100	6.760E+00	DenF	0.095	Emin	-0.834	0.24
90	4.526E+08	ETPa	-2.754	Emin	-0.460	0.35
80	3.366E-01	DenD	0.118	%Arc	0.011	0.33
70	3.654E-01	%Agr	-0.009	DenD	0.136	0.47
60	4.222E-01	ADre	0.024	DenD	0.102	0.61
50	5.001E-01	DenD	0.088	LonC	0.048	0.71
40	1.095	DenD	0.110	Emed	-0.069	0.78
35	1.281	DenD	0.096	Emed	-0.073	0.82
30	1.898	DenD	0.080	Emax	-0.098	0.84
20	2.197	DenD	-0.077	Pend	-0.077	0.52
10	2.331	%Bosq	-0.008	DenD	-0.070	0.45
5	2.996	LonC	-0.084	H75	0.047	0.57
1	2.884E+00	ADre	-0.095	Desn	0.113	0.71
0.5	2.925E+00	ADre	-0.114	Desn	0.148	0.75
0.1	4.027E+00	Perim	-0.259	Desn	0.217	0.77

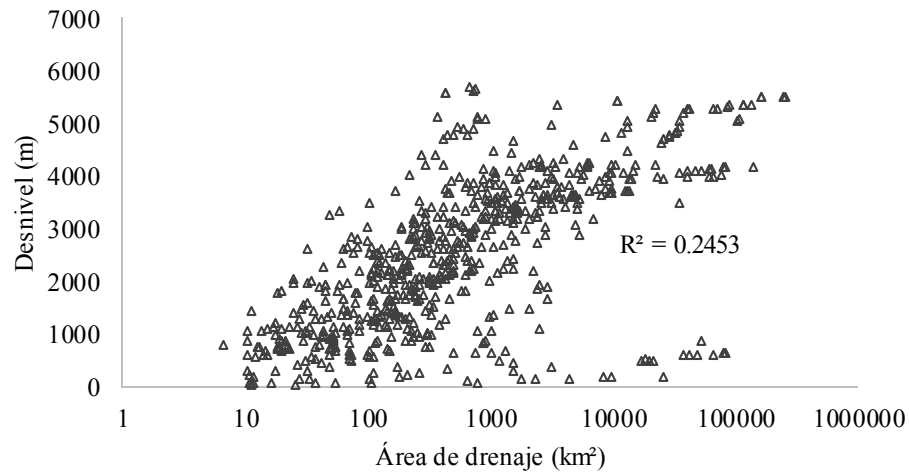
Ecuaciones de Regresión Región Pacífico

Caudal Q_p	β	X_1	α_1	X_2	α_2	R^2
100	3.977E-03	Komp	3.747	%Bosq	0.248	0.24
90	1.237E-01	%Urb	0.046	Pmed	0.262	0.51
80	1.936E-01	%Urb	0.036	Pmed	0.215	0.52
70	4.054E-01	ADre	0.044	%Bosq	0.024	0.53
60	5.162E-01	ADre	0.035	%Bosq	0.022	0.55
50	6.640E-01	ADre	0.024	%Bosq	0.017	0.54
40	8.635E-01	ADre	0.012	%Bosq	0.012	0.46
35	7.745E-01	%Bosq	0.009	DenD	0.088	0.33
30	6.173E-01	DenD	0.110	Pend	0.078	0.19
20	3.010E+03	ETPa	-1.578	Tmed	0.945	0.49
10	9.236E-01	Pmáx	0.400	Pmed	-0.349	0.50
5	3.143E+00	ADre	-0.041	%Bosq	-0.016	0.48
1	6.963E+00	Perim	-0.131	%Bosq	-0.030	0.51
0.5	1.008E+01	Perim	-0.172	%Bosq	-0.035	0.53
0.1	1.975E+01	Perim	-0.246	%Bosq	-0.043	0.54

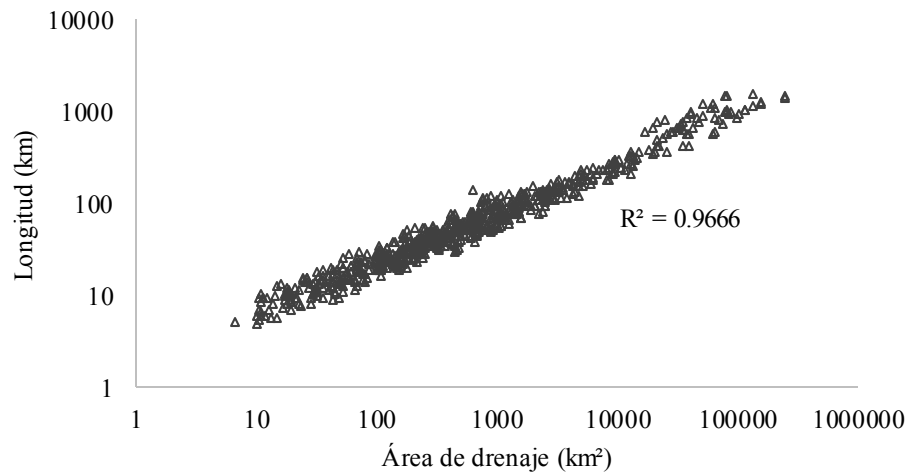
D. Dispersiones Atributos vs Adre



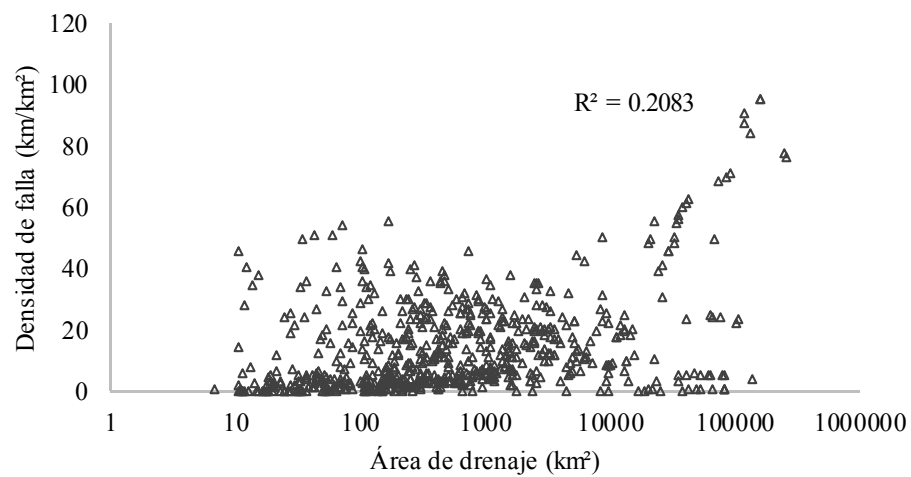
Desn Vs Adre



LonC Vs Adre



DenF Vs Adre



Bibliografía

- Álvarez, G., Hotait, N., & Sustaita, F. (2011). Identificación de regiones hidrológicas homogéneas mediante análisis multivariado. *Ingeniería Investigación y Tecnología*, *XIII*(3), 277–284.
- Anderson, T. (1958). *An introduction to multivariate statistical analysis*. New York, London, Sidney: John Wiley & Sons, Inc.
- Atieh, M., Taylor, G., Sattar, A., & Gharabaghi, B. (2017). Prediction of flow duration curves for ungauged basins. *Journal of Hydrology*, *545*, 383–394. <https://doi.org/10.1016/J.JHYDROL.2016.12.048>
- Barriendos, M., Coeur, D., Lang, M., Llast, M. C., Naulet, R., Lemaître, D., & Barrera, A. (2003). Stationarity analysis of historical flood series in France and Spain (14th–20th centuries). *Natural Hazards and Earth System Sciences*, (3), 583–592.
- Beven, K. J. (2012). *Rainfall-runoff modelling: the primer*. Wiley-Blackwell.
- Blöschl, G. (2005). Rainfall-Runoff Modeling of Ungauged Catchments. In *Encyclopedia of Hydrological Sciences*. Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/0470848944.hsa140>
- Blöschl, G. (2013). *Runoff prediction in ungauged basins: synthesis across processes, places and scales*. Cambridge University Press.
- Brown, A. E., Zhang, L., McMahon, T. A., Western, A. W., & Vertessy, R. A. (2005). A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. *Journal of Hydrology*, *310*(1–4), 28–61. <https://doi.org/10.1016/J.JHYDROL.2004.12.010>
- Burt, T. P., & Swank, W. T. (1992). Flow frequency responses to hardwood-to-grass conversion and subsequent succession. *Hydrological Processes*, *6*(2), 179–188. <https://doi.org/10.1002/hyp.3360060206>
- Cantor, D., & Ochoa, A. (2011). Señales de cambio climático en series de lluvia en Antioquia. In *IX Congreso Colombiano de Meteorología y Conferencia Internacional “200 años de la meteorología y de la climatología en América Latina”* (p. 11). Bogotá - Colombia.
- Castellarin, A. (2014). Regional prediction of flow-duration curves using a three-dimensional kriging. *Journal of Hydrology*, *513*, 179–191. <https://doi.org/10.1016/J.JHYDROL.2014.03.050>
- Castellarin, A., Camorani, G., & Brath, A. (2007). Predicting annual and long-term flow-

- duration curves in ungauged basins. *Advances in Water Resources*, 30(4), 937–953. <https://doi.org/10.1016/J.ADVWATRES.2006.08.006>
- Cataño-Álvarez, S., & Vélez Upegui, J. I. (2016). Aggregated conceptual model of sediment transport for mountain basins in Antioquia- Colombia. *Boletín de Ciencias de La Tierra*, (39), 38–48. <https://doi.org/10.15446/rbct.n39.52888>
- Chaves C., B., & Jaramillo R., A. (1998). Regionalización de la temperatura del aire en Colombia. *Repositorio Digital Del Centro Nacional de Investigación Del Café - Cenicafé*, 49(3), 224–230.
- Chávez, G., Jouravlev, A., & Dourojeanni, A. (2002). *Gestión del agua a nivel de cuencas: teoría y práctica*. Chile: Naciones Unidas, CEPAL, División de Recursos Naturales e Infraestructura.
- Comisión Nacional del Agua. (2006). Estadística del Agua en México, 2006. *CONCAGUA*, 200.
- Eeles, C., Robinson, M., & Ward, R. (1990). Experimental basins and environmental models. *Hydrologycal Research*, 44.
- Estupiñán, A. (2016). Estudio de la variabilidad espacio temporal de la precipitación en Colombia. *Bdigital.Unal.Edu.Co*, 118.
- Fennessey, N., & Vogel, R. M. (1990). Regional Flow-Duration Curves for Ungauged Sites in Massachusetts. *Journal of Water Resources Planning and Management*, 116(4), 530–549. [https://doi.org/10.1061/\(ASCE\)0733-9496\(1990\)116:4\(530\)](https://doi.org/10.1061/(ASCE)0733-9496(1990)116:4(530))
- Foster, H. A. (1933). Duration Curves. *Proceedings of the American Society of Civil Engineers*, 59(8), 1223–1246.
- Gallego, S., & Carvajal, L. F. (2017). Regionalización de curvas de duración de caudales en el Departamento de Antioquia-Colombia. *Revista EIA*, 14(27), 21–30. <https://doi.org/10.24050/REIA.V14I27.1158>
- García, M. C., Botero, A. P., Quiroga, F. A. B., & Robles, E. A. (2012). Variabilidad climática, cambio climático y el recurso hídrico en Colombia. *Revista de Ingeniería*, 0(36), 60–64. <https://doi.org/10.16924/riua.v0i36.136>
- García, P. L., Méndez, J. F., & Zárate, M. F. (2017). Delimitation of Colombia hydrologic regions. *Ingeniería y Desarrollo*, 35(1), 132–151. <https://doi.org/10.14482/inde.35.1.8946>
- Gilbert, R. O. (1987). *Statistical methods for environmental pollution monitoring*. Wiley.
- GWP. (2008). “Principios de gestión integrada de los recursos hídricos. Bases para el desarrollo de planes nacionales.
- Hamed, K. H., & Ramachandra, A. (1998). A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology*, 204(1–4), 182–196. [https://doi.org/10.1016/S0022-1694\(97\)00125-X](https://doi.org/10.1016/S0022-1694(97)00125-X)
- Hartigan, J. (1975). Clustering algorithms. Retrieved from <http://cds.cern.ch/record/105051>

- Hurtado, A. F., & Mesa, Ó. J. (2014). Reanalysis of monthly precipitation fields in Colombian territory. *DYNA*, 81(186), 251. <https://doi.org/10.15446/dyna.v81n186.40419>
- Instituto de Hidrología Meteorología y Estudios Ambientales IDEAM. (2014). Estudio Nacional del Agua 2014.
- IPCC. (2014). *Climate change 2014: synthesis report*.
- Jaramillo, A. (1989). Relación entre la evapotranspiración y los elementos climáticos . (Nota técnica). *Cenicafé*, 40(3).
- Krasovskaia, I., Gottscalk, L., Leblois, E., & Pacheco, A. (2006). Regionalization of flow duration curves. *Climate Variability and Change - Hydrological Impacts*, (308), 105–110.
- Li, M., Shao, Q., Zhang, L., & Chiew, F. H. S. (2010). A new regionalization approach and its application to predict flow duration curve in ungauged basins. *Journal of Hydrology*, 389(1–2), 137–145. <https://doi.org/10.1016/J.JHYDROL.2010.05.039>
- Mann, H. (1945). Nonparametric tests against trend. *Journal of the Econometric*, 13(3), 245–259.
- Mesa Sánchez, Ó. J., Vélez Upegui, J. I., Giraldo Osorio, J. D., Quevedo Tejada, D. I., Mesa Sánchez, Ó. J., Vélez Upegui, J. I., ... Quevedo Tejada, D. I. (2003). Regionalización de Características Medias de la Cuenca con Aplicación a Estimación de Caudales Máximos.
- Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. S. (2008). Stationarity is dead: Whither water management? *Science*, 319(5863), 571–575.
- Mohamoud, Y. M. (2008). Prediction of daily flow duration curves and streamflow for ungauged catchments using regional flow duration curves. *Hydrological Sciences Journal*, 53(4), 706–724. <https://doi.org/10.1623/hysj.53.4.706>
- Musiak, K., Inokuti, S., & Talahasi, Y. (1975). Dependence of low flow characteristics on basin geology in mountainous areas of Japan. *Proceeding of International Symposium of Hydrology, Tokyo, Japan, 1975*, 117, 147–156.
- Olden, J. D., Kennard, M. J., & Pusey, B. J. (2012). A framework for hydrologic classification with a review of methodologies and applications in ecohydrology. *Ecohydrology*, 5(4), 503–518. <https://doi.org/10.1002/eco.251>
- Pérez, C., Poveda, G., Mesa, O., & Carvajal, L. F. (1998). Evidencias de cambio climático en Colombia: tendencias y cambios de fase y amplitud de los ciclos anual y semianual. *Bulletin de l'Institut Français d'études Andines*, 27(3), 537–543.
- Prosdoci, I., Kjeldsen, T. R., & Svensson, C. (2013). Non-stationarity in annual and seasonal series of peak flow and precipitation in the UK. *Natural Hazards and Earth System Sciences Discussions*, 1(5), 5499–5544. <https://doi.org/10.5194/nhessd-1-5499-2013>

- Puertas, O., Carvajal, Y., & Quintero, M. (2011). Estudio de tendencias de la precipitación mensual en la cuenca alta-media del río Cauca, Colombia. *Revista Dyna*, 78(169), 112–120.
- Razavi, T., & Coulibaly, P. (2013). Streamflow Prediction in Ungauged Basins: Review of Regionalization Methods. *Journal of Hydrologic Engineering*, 18(8), 958–975. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000690](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000690)
- Rojo Abuín, J. M. (2007). Regresión Lineal Múltiple. *Institto de Economía y Gografía*, 2, 25.
- Salazar-Holguín, F. (2013). Zonificación Hidrográfica Preliminar de Colombia. *IDEM*.
- Salazar Oliveros, J. (2016). Una metodología para la estimación de curvas de duración de caudales (cdc) en cuencas no instrumentadas. Caso de aplicación para Colombia en los departamentos de Santander y Norte de Santander. *Repositorio Institucional - Universidad Nacional de Colombia*.
- Sauquet, E., & Catalogne, C. (2011). Comparison of catchment grouping methods for flow duration curve estimation at ungauged sites in France. *Hydrology and Earth System Sciences*, (15), 2421–2435.
- Sawicz, K., Wagener, T., Sivapalan, M., Torch, P. A., & Carrillo, G. (2011). Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrology and Earth System Sciences*, (15), 2895–2911.
- Sen, P. K. (1968). Estimates of the Regression Coefficient Based on Kendall's Tau. *Journal of the American Statistical Association*, 63(324), 1379–1389. <https://doi.org/10.1080/01621459.1968.10480934>
- Sivapalan, M. (2003). Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological Processes*, 17(15), 3163–3170. <https://doi.org/10.1002/hyp.5155>
- Steel, R. G. D., & Torrie, J. H. (1960). *Principles and Procedures of Statistics*. New York, Toronto, London: McGRAW-HILL BOOK COMPANY, INC.
- UPME-PPUJ. (2015). *Atlas potencial hidroenergético de Colombia*.
- Urán, J. D. (2016). Cambios en los valores extremos de variables climáticas en Colombia asociados a cambio climático. *Repositorio Institucional - Universidad Nacional de Colombia*, 121.
- Vaze, J., Post, D. A., Chiew, F. H. S., Perraud, J.-M., Viney, N. R., & Teng, J. (2010). Climate non-stationarity – Validity of calibrated rainfall–runoff models for use in climate change studies. *Journal of Hydrology*, 394(3–4), 447–457. <https://doi.org/10.1016/J.JHYDROL.2010.09.018>
- Villarini, G., Serinaldi, F., Smith, J. A., & Krajewski, W. F. (2009). On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resources Research*, 45(8). <https://doi.org/10.1029/2008WR007645>
- Vogel, R. M., & Fennessey, N. M. (1994). Flow-Duration Curves. I: New Interpretation and

- Confidence Intervals. *Journal of Water Resources Planning and Management*, 120(4), 485–504. [https://doi.org/10.1061/\(ASCE\)0733-9496\(1994\)120:4\(485\)](https://doi.org/10.1061/(ASCE)0733-9496(1994)120:4(485))
- Waseem, M., Ajmal, M., & Kim, T.-W. (2016). Improving the flow duration curve predictability at ungauged sites using a constrained hydrologic regression technique. *KSCE Journal of Civil Engineering*, 20(7), 3012–3021. <https://doi.org/10.1007/s12205-016-0038-z>
- Westra, S., & Sisson, S. A. (2011). Detection of non-stationarity in precipitation extremes using a max-stable process model. *Journal of Hydrology*, 406(1–2), 119–128. <https://doi.org/10.1016/J.JHYDROL.2011.06.014>
- Wilks, D. S. (2011). Cluster Analysis. *International Geophysics*, 100, 603–616. <https://doi.org/10.1016/B978-0-12-385022-5.00015-4>
- WMO. (1995). *INFOHYDROManal, Second edition, Operational Hydrology Report*. Geneva.